Folding and replication in complex dynamic molecular networks

Liu, Bin

DOI:
10.33612/diss.99784510

Chapter 2 Complex Molecules that Fold like Proteins Can Emerge Spontaneously

Folding can bestow macromolecules with various properties, as evident from nature's proteins. Until now complex folded molecules are either the product of evolution or of an elaborate process of design and synthesis. We now show that molecules, that fold in a well-defined architecture of substantial complexity, can emerge autonomously and selectively from a simple precursor. Specifically, we have identified a self-synthesizing macrocyclic foldamer with a complex and unprecedented secondary and tertiary structure, that constructs itself highly selectively from 15 identical peptide-nucleobase subunits, using a dynamic combinatorial chemistry approach. Folding of the structure drives its synthesis in 95% yield from a mixture of interconverting molecules of different ring sizes in a one-step process. Single crystal X-ray crystallography and NMR reveals a folding pattern based on an intricate network of noncovalent interactions involving residues space apart widely in the linear sequence. These results establish dynamic combinatorial chemistry as a powerful approach to developing synthetic molecules with folding motifs of a complexity that goes well beyond that accessible with current design approaches. The fact that such molecules can form autonomously implies that may have played a role in the origin of life at earlier stages than previously thought possible.

This chapter has been published:

Liu, B.; Pappas, C. G.; Zangrando, E.; Demitri, N.; Chmielewski, P. J.; Otto. S. *J. Am. Chem. Soc.***2019**, *141*, 1685.

Chapter 2

2.1 Introduction

The chemistry of life relies on biopolymers (proteins, nucleic acids) folding into specific conformations that dictate their properties. It is generally believed that the complex folded structures encountered in biology are the result of millions of years of evolution. Much of the research on synthetic foldamers[1-13] is driven by desire to bypass evolution, go beyond the constraints of using only nature's building blocks and directly access structures that fold like proteins, but are based on completely synthetic structures. The ultimate goal is achieving new and sophisticated functions that require the molecular complexity of extended folded structures. This goal is still largely out of reach and foldamers able to exhibit specific function[14-16] have remained rare, due to the huge challenge of obtaining new modes of folding in designed proteins[17,18], molecules that mimic peptides or nucleic acids[5,8,9], and in completely abiotic molecules[1-4,6,7,10-13].

The approach taken to accessing new synthetic foldamers has until now relied almost exclusively on design, followed by multi-step synthesis. An impressive new range of backbones have been developed that fold into a variety well-defined architectures, including secondary structures such as helices, and sheets[19]. Yet, the design approach tends to be based on relatively simple and small-range assembly motifs, primarily driven by interactions between residues close to each other in the oligomeric chain of monomer units. Foldamers that rely on long-range interactions (between residues further apart in the oligomeric chain, as observed in the folding of proteins and nucleic acids) have remained difficult to access due to a lack of reliable design rules. Hence, alternative approaches are needed for accessing fundamentally new classes of foldamers that rely on long-range interactions.

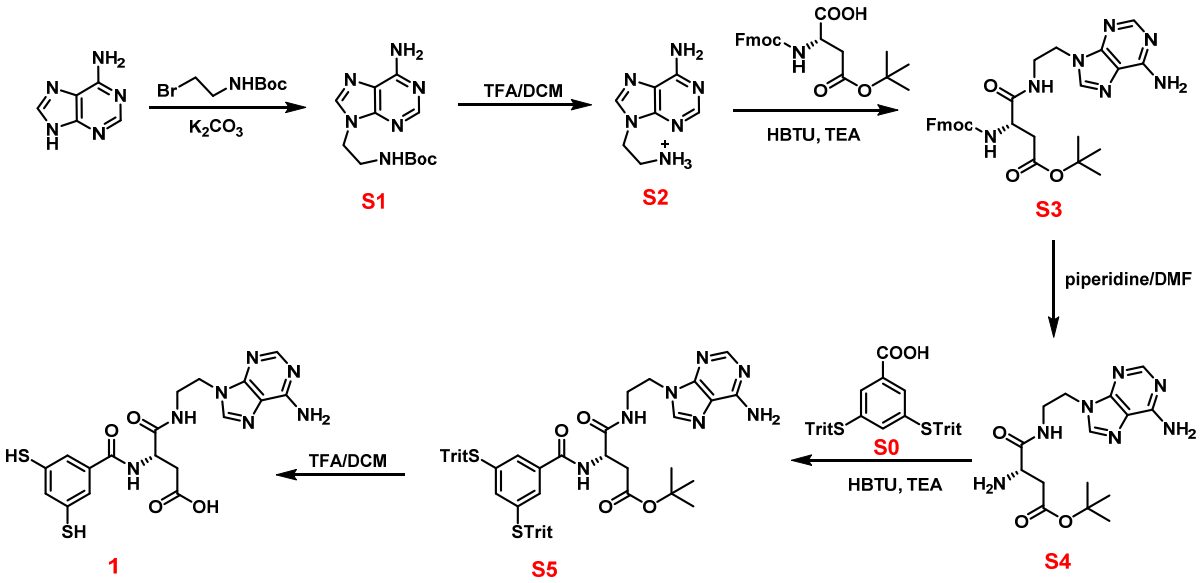
Dynamic combinatorial chemistry[20,21] has been suggested as a useful selection tool for accessing such new folded structures. In brief, in a dynamic combinatorial library building blocks react with each other to give rise to a mixture of oligomeric compounds that continuously exchange these building blocks between them. When a specific library member is able to form efficient intramolecular noncovalent interactions, inducing it to fold, this compound should be more stable than other library members that are unable to engage in such interactions. Hence, the library composition should shift in favor of the foldamer. Indeed, several groups have reported folding-driven changes in library compositions[22-29]. However, until now this approach has failed to deliver fundamentally new folding motifs. This lack of success is surprising, given that dynamic combinatorial methods have proven successful in the discovery of new and often unexpected host-guest systems[30,31], interlocked structures[31,32] and self-replicating molecules[33]; all systems also rely on non-covalent interactions as the prime selection criterion.

We now report results that should prompt a revival of the combinatorial approach to foldamers. We discovered a new macrocyclic foldamer with a complex secondary and tertiary structure, constructed out of 15 identical peptide-nucleobase subunits. Folding of the structure drives its synthesis in 95% yield from a mixture of interconverting molecules of different ring size in a one-step process. Single crystal X-ray crystallography and NMR reveals a new folding pattern based on an intricate network of hydrogen bonds and π-stacking interactions featuring long-range interactions.

2.2 Results and discussion

2.2.1 Design and synthesis of building blocks

We designed chimeric building block **1** that contains an amino-acid and a nucleobase subunit; both key structural units involved in the folding of proteins and nucleic acids, respectively. We reasoned that the presence of these two types of monomer units, that are central in nature's folded macromolecules, should maximize the chances of accessing foldamers in our dynamic libraries. At the same time, and unlike nature, having both monomer units present within a single building should give rise to fundamentally new foldamers. Building block **1** was synthesized in six steps as outlined in **Scheme 2.1**. The building block is equipped with two thiol groups, which, upon exposure to atmospheric oxygen, oxidize to give rise to an equilibrium mixture of different macrocyclic disulfides, which interconvert through thiol-disulfide exchange.



Scheme 2.1. Synthesis of building block **1**.

2.2.2 Formation of complex folded structures from DCLs

Ordinarily DCLs of building blocks such as **1** are dominated by small trimeric and tetrameric macrocycles (the dimer is too strained to form in significant quantities). Indeed, a small DCL made

from a 50 μM solution of building block **1** in borate buffer (50 mM, pH 8.2) was dominated by the cyclic tetramer (**Figure 2.1** trace 1; composition monitored by ultra-performance liquid chromatography - UPLC). However, repeating the same experiment at a higher building blocks concentration (500 μM) led to the emergence of an unusually large macrocycle, consisting of 15 subunits of **1**, as evident from mass spectrometric analysis of the relevant peak in the UPLC chromatogram (**Figure 2.21**). Increasing the concentration of the building block or addition of NaCl (1.0 M) to the mixture enhanced the formation of the 15mer, accounting for up to 95% of the library material (**Figure 2.1** trace 2). Experiments with other salts revealed similar effects (**Figure S2.4**), indicating that the exact nature of the salt is not important and that the formation of the 15mer benefits from a high ionic strength, suggesting that the salt acts to reduce charge repulsion within the 15mer. Experiments in the absence of salt but in the presence of 50% acetone (trace 5 in **Figure 2.1**) or other cosolvents (**Figure S2.5**) gave only negligible amounts of 15mer, suggesting that hydrophobic interactions are important in stabilizing this compound. These observations, together with the fact that the 15mer is only poorly retained on the UPLC column, suggest that the compound adopts a folded structure. This hypothesis was confirmed when we characterized its structure using tandem mass spectrometry, circular dichroism (CD) spectroscopy, X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy.

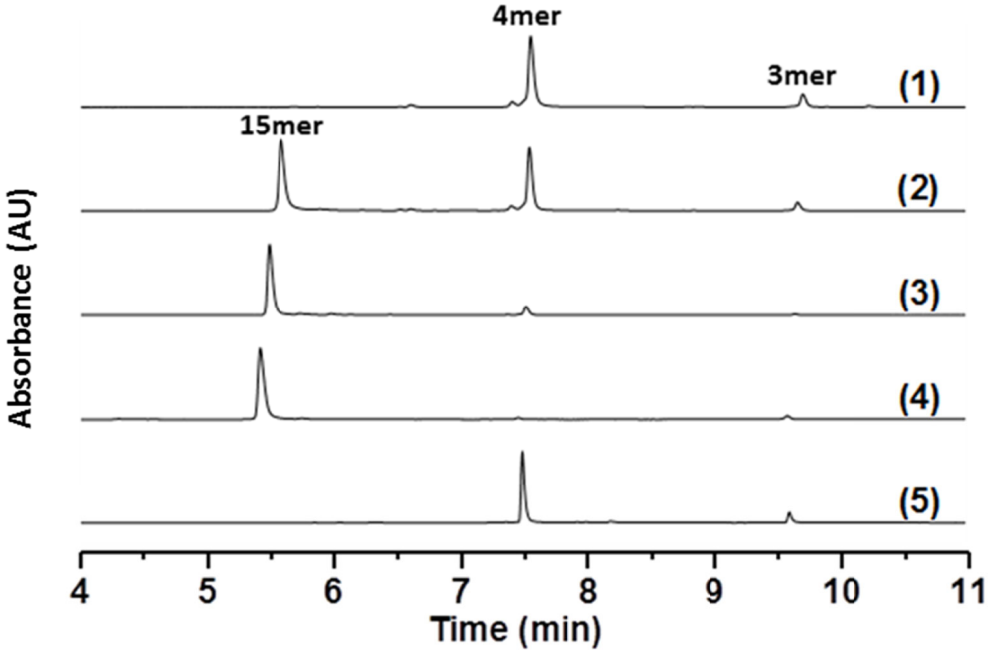


Figure 2.1. UPLC traces (absorption at 254 nm) showing library compositions after 14 days of stirring (in 50 mM borate buffer, pH = 8.2) at a building block concentration of: (**1**) 0.050 mM; (**2**) 0.50 mM; (**3**) 2.0 mM; (**4**) 0.50 mM in the presence of 1.0 M NaCl; (**5**) 0.50 mM in the presence of 50% acetone.

2.2.3 Structure characterization of foldamer

The 15mer was isolated by preparative high-performance liquid chromatography (HPLC) (isolated yield: 90%) and then analyzed by means of matrix-assisted laser desorption ionization time-of-flight tandem mass spectrometry (MALDI TOF/TOF). The results (**Figure 2.2**) show that the 15mer fragments into smaller components, ranging from tetramer up to 12mer, indicating that the 15mer is a single macrocycle and not a system of interlocked rings (a catenane)[31].

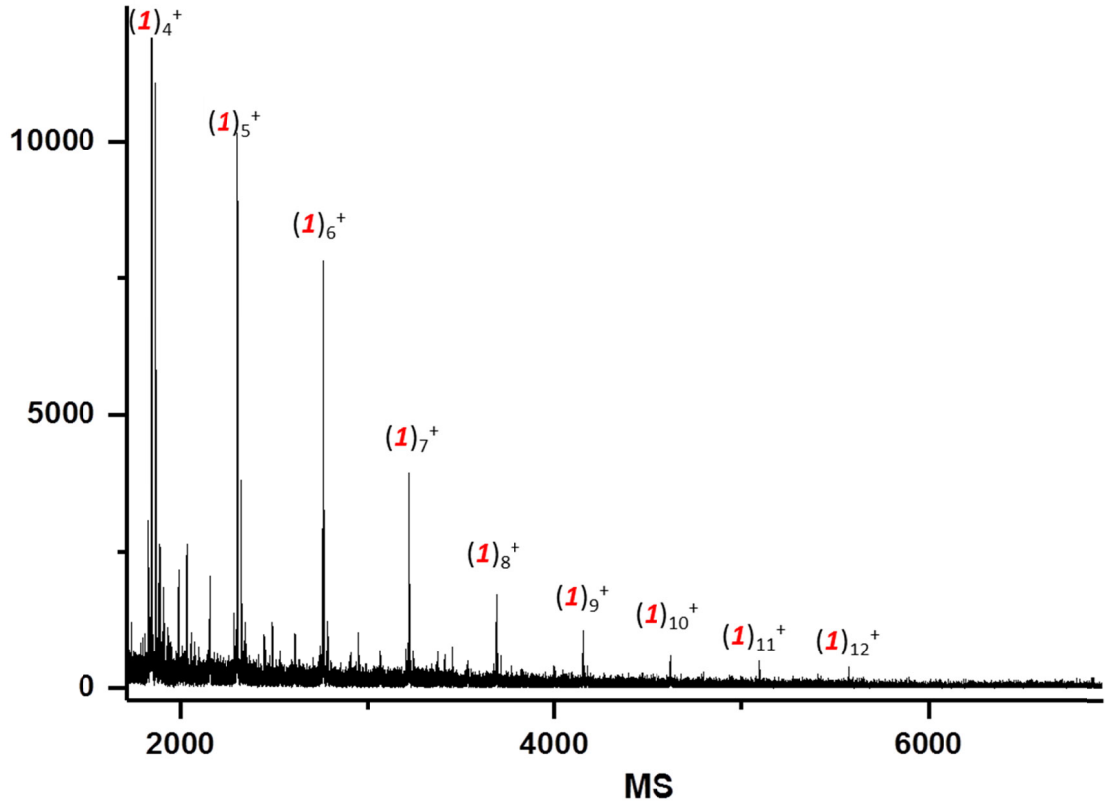


Figure 2.2. TOF/TOF analysis of foldamer $\mathbf{1}_{15}$.

The CD spectrum of the 15mer showed an intense positive band at 260 nm, which can be attributed to the absorption of the aromatic dithiol[34] and the terminal adenine. Importantly, the CD signals of the 15mer are dramatically enhanced compared to those of monomer **1** (Figure 2.3), which suggests that the aromatic rings reside in a well-defined chiral environment, even though these rings are relatively remote from the chiral center in building block **1**, located on the amino-acid residue.

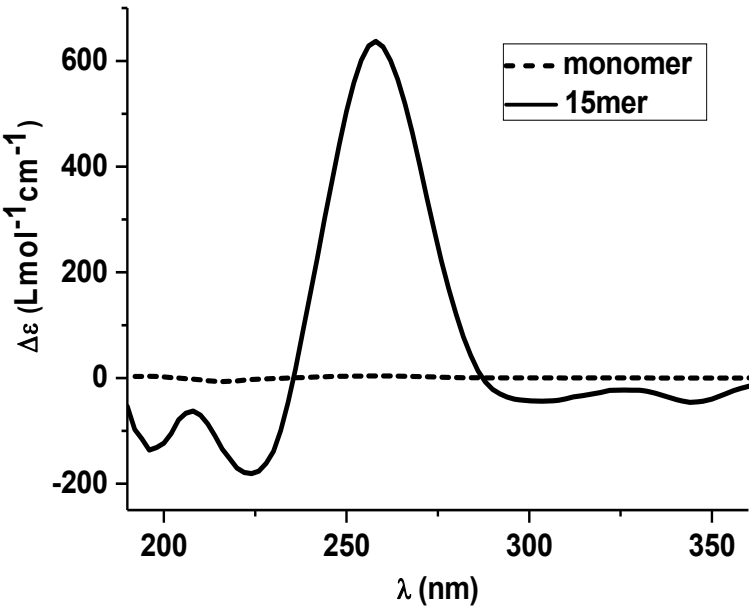


Figure 2.3. CD spectrum of monomer **1** and 15mer in water at 298 K.

Detailed insights into the structure of 15mer were obtained from single crystal X-ray diffraction data. Crystals of the 15mer were prepared by slow diffusion of acetone into an aqueous solution of the 15mer. The crystal structure (**Figure 2.4**) confirms that the 15mer is a single giant macrocycle[35] linked together by 15 disulfide bonds to give a 75-atom ring. **Figure 2.4a** shows the extended conformation of this ring for clarity, which, in reality, is collapsed into a compact but intricately folded structure (**Figure 2.4e** shows the 75-atom ring in its highly twisted conformation). Overall, the structure is characterized by a hydrophobic core and presents its hydrophilic groups on its surface, similar to what is observed in folded proteins. The most notable structural motif is the stacking of aromatic rings. Five stacks of three phenyl rings can be identified (shown in **Figure 2.4f**), capped with an adenine ring at the top and bottom (except where these adenines are recruited for crystal packing) as shown for one of these stacks in **Figure 2.4g**. The distances between the three phenyl rings in these stacks are in the range of 3.43-3.45 Å, while the distances between the phenyl and adenine rings are somewhat larger (3.46-3.49 Å), but all in the range typical for π-stacking[36]. Interestingly, the rings that end up in the same stack are spaced far apart in the extended structure; as indicated in **Figure 2.4a**; a stack is composed of the phenyl rings from the *i*, *i*+2 and *i*+4 residues, while the capping adenines belong to the *i*-3 and the *i*+7 residues. Except for the adenines recruited for crystal packing (disrupting the otherwise 5-fold symmetry of the structure), this arrangement gets repeated five times in an interdigitated fashion (the second stack is indicated by a dotted line in **Figure 2.5a**, the other stacks are not shown for clarity). The five stacks of rings are arranged in a tiled fashion as

shown in **Figure 2.4f**. Thus, not only secondary structure elements can be identified (corresponding to the stacks of rings) but since these stacks adopt well-defined orientations with respect to each other, also a tertiary structure is present, which has only very recently been achieved in designed foldamers[37]. Apart from π-stacking interactions also hydrogen bonds play an important role in stabilizing this folded structure. Between the building blocks that constitute the π-stack (residues *i*, *i*+2 and *i*+4), five hydrogen bonds are observed between the NH and CO groups (**Figure 2.4h**).

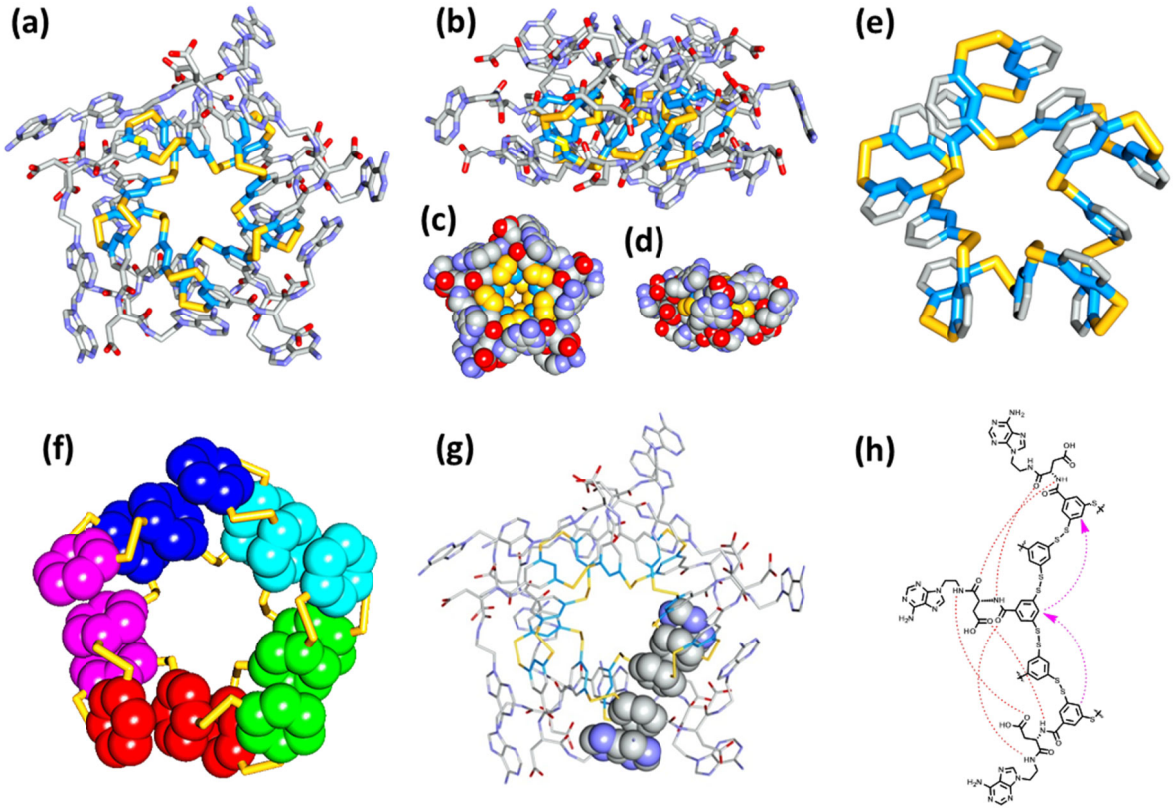


Figure 2.4. X-ray crystal structure of the 15mer. (**a**) Top view of the central cavity of the macrocycle; (**b**) Side view; (**c**) Top and (**d**), side view of the 15mer in space filling representation; (**e**) The ring of disulfide bonds connecting the phenyl rings; (**f**) Core part of the foldamer, showing five stacks of three phenyl rings connected by disulfide bonds; (**g**) Top view of the 15mer highlighting one stack of three core phenyl rings and two adenines on the top and bottom of the stack; (**h**) Set of intermolecular hydrogen bonds formed between the three building blocks that constitute the stack of three phenyl rings. Solvent molecules, hydrogen atoms and disorders are omitted for clarity. C atoms are shown in gray, N in purple, O in red, and S in yellow; except in panel f the C atoms of the macrocycle core are shown in light blue.

The solution-phase ^{1}H-NMR spectrum (**Figure 2.5c**) of the isolated 15mer (D_2O, 298K) is consistent with the X-ray crystal structure. The presence of sharp signals is in agreement with a highly ordered compact structure. In the spectrum the monomeric unit **1** appears in three separate sets of signals, indicative of a C_5 symmetry for the 15mer in solution. The signals of the protons of the phenyl rings

are distributed over a wide range of chemical shifts (from 5.8 to 7.6 ppm). The large upfield shift of two of these (to 6.6 and 5.8 ppm respectively, versus 7.2 and 7.3 ppm in monomer **1**), indicates that some of them are adjacent to the face of an aromatic ring. 2D-NMR studies enabled the complete assignment of the spectrum of the 15mer as well as the assignment of several through-space interactions that are consistent with the x-ray crystal structure (**Figure 2.5b**). These data indicate that the 15mer adopts a folded structure also in solution.

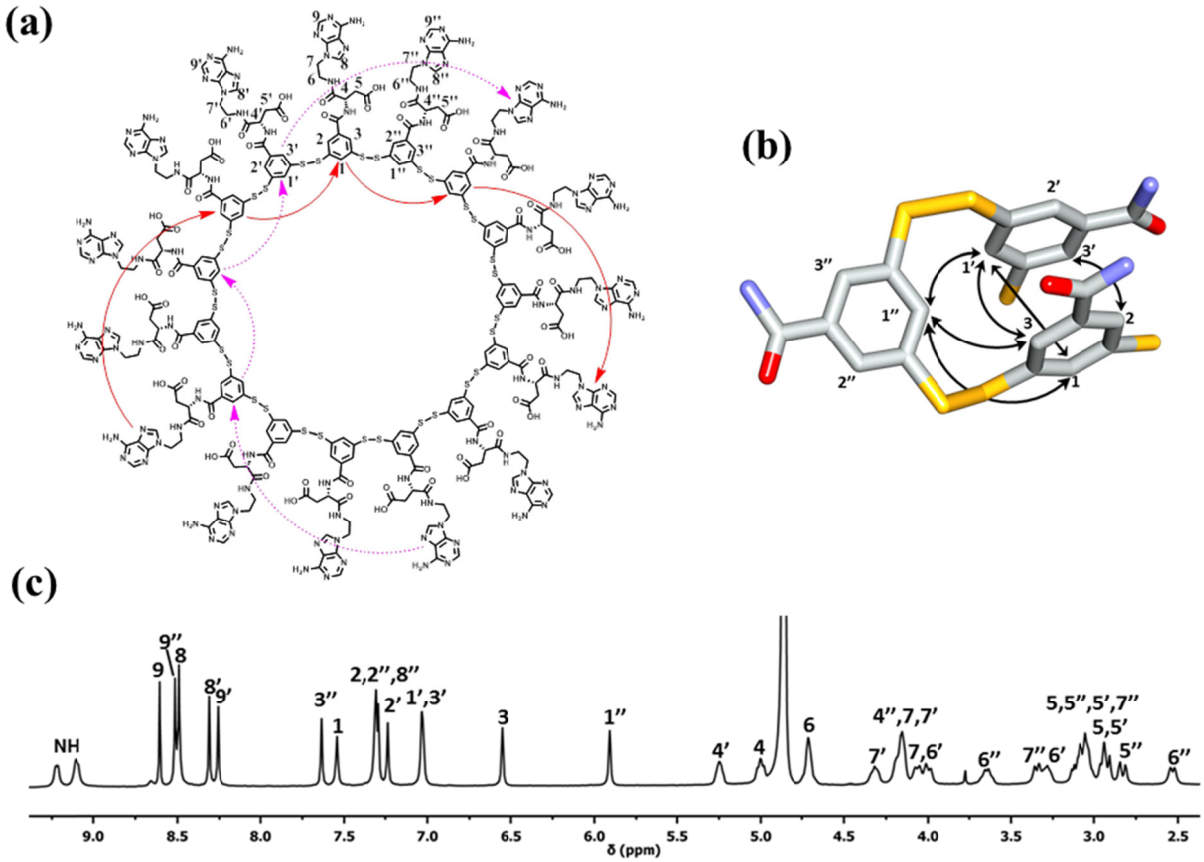


Figure 2.5. NMR spectra studies of foldamer. (**a**) Chemical structure of 15mer (Arrows indicate two sets of π-stacks observed between phenyls and adenines. The remaining three stacks are not shown for clarity); (**b**) One-fifth core part of 15mer with observed NOEs; (**c**) ^{1}H-NMR spectrum of the 15mer (500 MHz, D_2O, 298K), with signal assignments corresponding to the labeling shown in (**a**).

2.2.4 Unfolding and refolding

We then attempted to unfold the 15mer. Temperature dependent ^{1}H-NMR experiments on aqueous solutions of 15mer were performed, but no significant spectral changes were observed, even at temperatures up to 353 K (**Figure S2.9**). Acetone was added (which disrupts the process of formation of the 15mer; vide supra) but the NMR spectra remained essentially unchanged (**Figure S2.10**), indicating that the 15mer, once formed, does not readily unfold. Unfolding was eventually achieved

by dissolving the 15mer in DMF-d_7. At room temperature sharp signals corresponding to the folded 15mer were still observed (**Figure S2.11**) but these broadened significantly when the temperature was increased to 363 K, suggesting (at least partial) unfolding. At 373 K one set of relatively sharp signals was observed indicating a 15-fold symmetry of the system, and thus, complete unfolding of the macrocycle. The original spectrum was retained upon cooling the sample down to room temperature. The unfolding-refolding process was further analyzed by variable temperature CD spectroscopy in DMF (**Figure S2.13**). Reversible attenuation and enhancement of the CD signals were observed by heating up and cooling down the DMF solution (**Figure 2.6**).

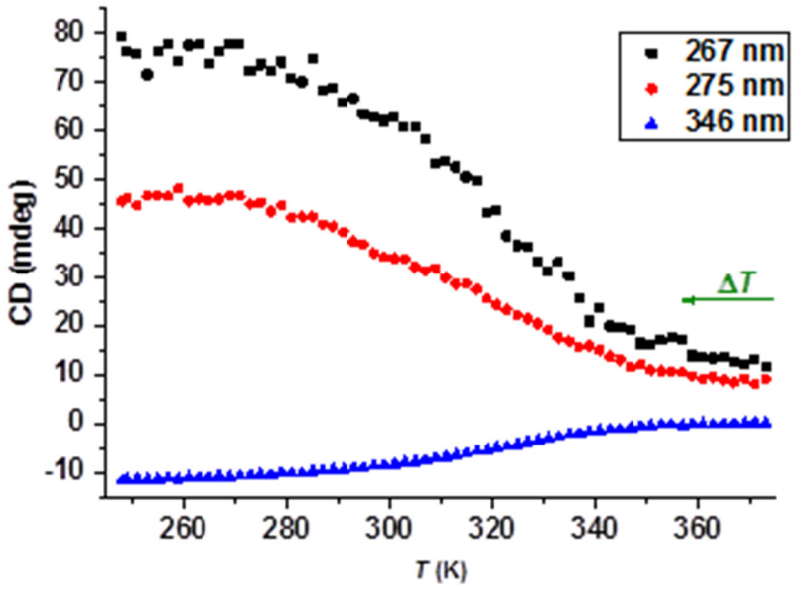


Figure 2.6. Changes in the Cotton effect intensities at specified wavelengths in the CD spectra of a DMF solution of the 15mer observed upon cooling from 373 K to 248 K.

2.2.5 Building block modification

Finally, we investigated the extent to which foldamer formation depends on the structure of the building block. Foldamer formation was critically dependent on the presence of the nucleobase, but not specific for adenine, as replacement of adenine by guanine also gave 15mer in 95% yield (**Figure 2.7b**), while without any nucleobase only cyclic trimers and tetramers were detected (**Figure 2.7d**).

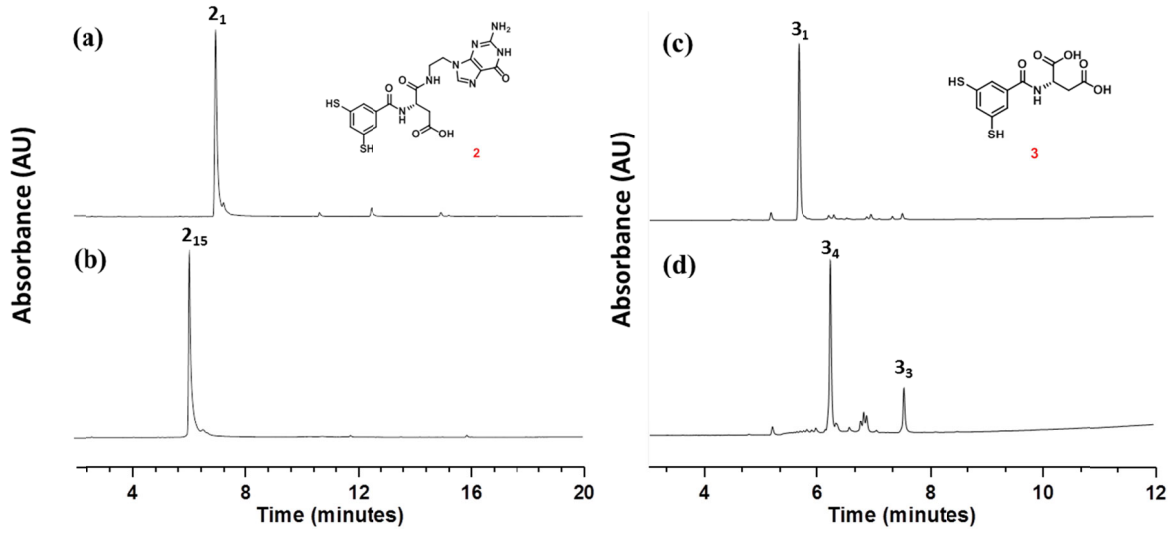


Figure 2.7. UPLC analyses (absorption at 254 nm) of the product distribution of libraries made from building blocks **2** and **3** in borate buffer (pH=8.2, 50 mM) at 1.0 mM concentration: (**a**) and (**c**) immediately after dissolving building block **2** and **3**; (**b**) and (**d**) after stirring for 11 days.

2.3. Conclusion

In conclusion, these results establish dynamic combinatorial chemistry as a promising tool for identifying new foldamers with unprecedented folds of substantial structural complexity that are not readily predictable. The dynamic combinatorial approach complements existing design approaches by allowing access to structures that feature interactions between residues spaced apart widely in the linear sequence; such structures remain difficult to access by design. This method allows selective access to oligomers of precisely defined length, in remarkably high yield and with minimal synthetic effort. Our findings are also relevant in the context of the origin of life, as they proof that folded structures of considerable complexity can emerge selectively and spontaneously from a mixture of interconverting molecules, simply as a result of their thermodynamic stability. Folding drives the synthesis of the molecule. Such folded structures could therefore have played a role in prebiotic chemistry that led to the first forms of life. Where complex folded structures were, until now, generally considered to be products of evolution, our results show that such structures do not require an evolutionary process, but can also form autonomously. Thus, they may feature much earlier in scenarios of the origins of life than previously thought possible.

2.4. Experimental section

2.4.1 General methods

All chemicals, unless otherwise stated, were purchased from Sigma-Aldrich and used as received. Acetonitrile (ULC-MS grade), water (ULC-MS grade) and trifluoroacetic acid (HPLC grade) were purchased from Biosolve BV. Anhydrous solvents used in synthesis were freshly collected from a dry solvent purification system prior to use. Flash column chromatography was performed on a Reveleris® X2 Flash Chromatography System (Grace Davison Discovery Sciences, Deerfield IL) on normal or reverse phase silica cartridges. NMR spectra were recorded on 400, 500 and 600 MHz, cryo-NMR spectrometers, locked on deuterated solvents and referenced to the solvent peak. HRMS spectra were recorded on a LTQ Orbitrap XL instrument in ESI mode.

Buffer preparation

Borate buffer (50 mM, pH = 8.0) was prepared by dissolving sodium tetraborate (3.841 g, $Na_2B_4O_7.10H_2O$) in 200 mL doubly distilled water. Then the pH was adjusted by concentrated HCl to pH 8.0.

Library preparation and sampling

Building blocks were dissolved in borate buffer (50 mM, pH 8.2) to prepare a library. All libraries were set up in an HPLC vial (12×32 mm) with a Teflon-coated screw cap. All HPLC vials were equipped with a cylindrical stirrer bar (2×5 mm, Teflon coated, purchased from VWR) and were stirred at 1200 rpm using an IKA RCT basic hot plate stirrer. All experiments were performed at ambient conditions.

UPLC and UPLC-MS analysis

UPLC analyses were performed on a Waters Acquity H-class system equipped with a PDA detector, at a detection wavelength of 254 nm. Samples were injected on an Phenomenex Aeris Peptides 1.7 µm (150 × 2.1 mm) column, purchased from Phenomenex, using ULC-MS grade water (eluent A) and ULC-MS grade acetonitrile (eluent B), containing 0.1 V/V % TFA as modifier. A flow rate of 0.3 mL/min and a column temperature of 35 °C were applied.

UPLC-MS analyses were performed using a Waters Acquity UPLC H-class system coupled to a Waters Xevo-G2 TOF. The mass spectrometer was operated in positive electrospray ionization mode with the following ionization parameters: capillary voltage: 3 kV; sampling cone voltage: 20 V; extraction cone voltage: 4 V; source gas temperature: 120 °C; desolvation gas temperature: 450 °C; cone gas flow (nitrogen): 1 L/h; desolvation gas flow (nitrogen): 800 L/h.

Chapter 2

Circular Dichroism

Spectra were recorded at room temperature by using a JASCO J715 spectrophotometer and HELMA quartz cuvettes with a path length of 1 mm. All spectra were recorded at room temperature from 190 nm to 300 nm, with 2 nm step interval and 3 scans with a speed of 200 nm/min. Solvent spectra were subtracted from all spectra. Samples were diluted to a concentration of 0.15 mM with respect to building block concentration.

TOF/TOF analysis

Sample volumes of 1 µL were applied to a MALDI target plate (stainless steel, polished), and mixed on the plate with 1 µL of matrix solution, containing 5 mg/mL α-cyanohydroxycinnamic acid in water/acetonitrile 50/50 v/v with 0.1% trifluoroacetic acid. After drying of the spots, positive reflector mode MALDI-TOF spectra were recorded between m/z 825 and 6000 with an UltrafleXtreme MALDI-TOF/TOF mass spectrometer (Bruker Daltonics, Bremen, Germany) operated under Bruker flexControl software (version 3.4); 8000 shots were acquired randomly over the spot.

UPLC methods

Methods for the analysis of DCLs made from building block **1**:

Table S2.1. Method 1:

t / min	% B
0	10
10	40
12	90
13	90
14.5	10
17	10

Table S2.2. Method 2:

t / min	% B
0	10
20	40
22	90
27	90
28	10
30	10

Table S2.3. Method for the analysis of DCLs made from building block **2**:

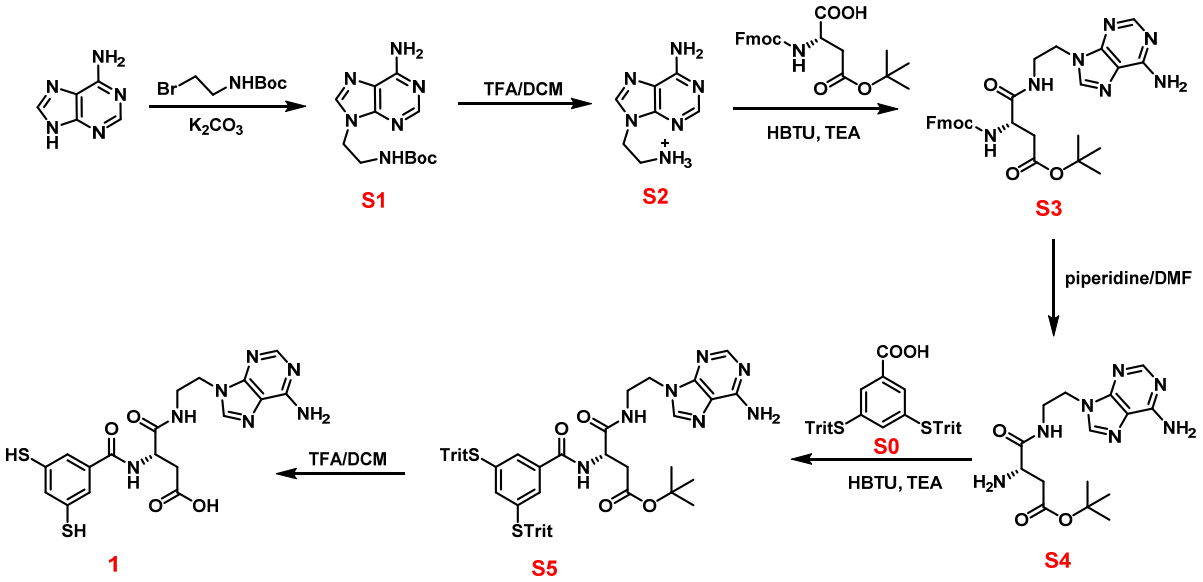
t / min	% B
0	10
2	15
8	20
20	40
22	90
27	90
28	10
30	10

Table S2.4. Method for the analysis of DCLs made from building block **3**:

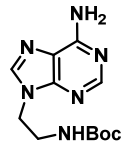
t / min	% B
0	10
1	30
11	45
12	90
13	90
14.5	10
17	10

2.4.2 Synthesis and characterization of building blocks

Synthesis of building block **1**



S1

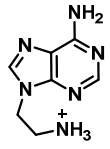


tert-Butyl (2-bromoethyl)carbamate (4.2 g, 15 mmol) was added to a dry dimethylformamide (DMF, 120 mL) solution of adenine (2.0 g, 15 mmol) and K_2CO_3 (5.0 g, 36 mmol) and the reaction mixture was stirred at room temperature for 48 hours. Then the solvent was evaporated under vacuum and the crude mixture was purified by flash column chromatography (SiO_2, 0-10% methanol in DCM), followed by removal of solvent to afford product **S1** as a colorless solid.

Yield = 34%, 1.4 g.

^{1}H NMR: (400 MHz, CD_3OD, 298K) δ_H = 1.31 (s, 9H, -Boc), 3.47-3.48 (m, 2H, -CH_2), 4.28-4.29 (m, 2H, -CH_2), 8.02 (s, 1H, adenine H), 8.19 (s, 1H, adenine H). ^{13}C NMR (101 MHz, CD_3OD, 298K) δ_C = 29.9, 42.2, 46.1, 81.5, 121.2, 144.2, 152.1, 154.8, 158.5, 159.4. HRESI-MS calc. for $C_{12}H_{19}N_6O_2^+$ 279.1564, found 279.1593.

S2

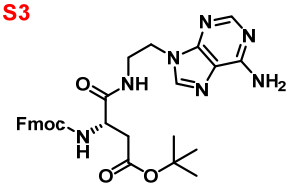


S1 (0.5 g, 1.8 mmol) was dissolved in 10 mL 50% TFA/DCM under a nitrogen atmosphere and the mixture was stirred overnight. After that, the solvents were evaporated under reduced pressure and the crude product was dissolved in methanol and precipitated by addition of diethylether. Colorless solid powder was collected by filtration.

Yield = 98%, 0.49g.

^{1}H NMR: (400 MHz, DMSO-d6, 298K) δ_H = 3.34-3.35 (m, -CH_2, 2H), 4.45 (t, *J* = 6.4 Hz, -CH_2, 2H), 8.09 (br, -NH_3, 3H), 8.35 (s, 1H, adenine H), 8.43 (s, 1H, adenine H), 8.82 (br, -NH_2, 2H). ^{13}C NMR (101 MHz, DMSO-d6, 298K) δ_C = 38.6, 41.9, 118.8, 114.1, 146.3, 149.6, 151.5. HRESI-MS calc. for $C_7H_{11}N_6^+$ 179.1040, found 179.1042.

S3



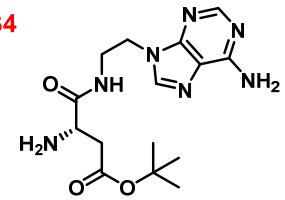
L-Fmoc-aspartic acid alpha-*t*-butylester (411 mg, 1.00 mmol), 2-(*1H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU, 379 mg, 1.00 mmol), **S2** (292 mg, 1.00 mmol) and triethylamine (0.30 mL, 2.15 mmol) were dissolved in 10 mL dry acetonitrile and the mixture was stirred at room temperature overnight under a nitrogen atmosphere. Solvent was evaporated under vacuum and the crude mixture was purified by flash column chromatography (SiO_2, 0-10% methanol in DCM), followed by removal of solvent to afford **S3** as a white powder.

Yield = 58%, 330 mg.

^{1}H NMR: (400 MHz, DMSO-d6, 298K) δ_H = 1.33 (s, 9H, -$(CH_3)_3$), 2.38 (dd, J_1 = 16 Hz, J_2 = 9.6 Hz, 1H, -CH_2), 2.59 (dd, J_1 = 16 Hz, J_2 = 4.8 Hz, 1H, -CH_2), 3.38-3.51 (m, 2H, -CH_2), 4.16-4.33 (m, 6H), 7.18 (s, 2H, NH_2), 7.30 (t, *J* = 7.6 Hz, 2H, ArH), 7.40 (t, *J* = 7.6 Hz, 2H, ArH), 7.61 (d, *J* = 8.4 Hz, 1H, NH), 7.70 (t, *J* = 6.4 Hz, 2H, ArH), 7.86 (d, *J* = 7.6 Hz, 2H, ArH), 8.00 (s, 1H, adenine H), 8.12 (s, 1H, adenine H), 8.16 (t, *J* = 6.0 Hz, 1H, -CONH). ^{13}C NMR (101 MHz, DMSO-d6, 298K) δ_C = 27.8, 37.5, 42.6, 46.1, 46.7, 51.6, 65.8, 80.2, 118.8, 120.2, 125.34, 125.36, 127.1, 127.8, 140.8, 141.1, 143.8, 143.9, 149.7, 152.4, 155.8, 156.0, 169.5, 170.8. HRESI-MS calc. for $C_{30}H_{34}N_7O_5^+$ 572.2616, found 572.2608.

S3 (571 mg, 1.00 mmol) was dissolved in 20 mL 20% piperdine in DMF (20:80/piperidine:DMF) and stirred for 0.5 h at room temperature. Then the solvents were evaporated under vacuum and the crude powder was washed with hexane (30 mL x 3) to afford derivative **S4** as a slightly yellow powder.

S4

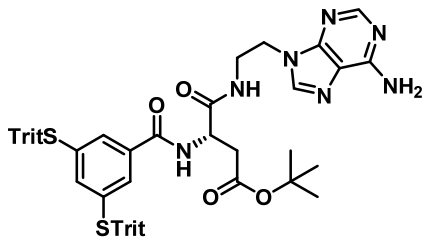


Yield = 63%, 220 mg.

^{1}H NMR: (400 MHz, CD_3OD, 298K) δ_H = 1.43 (s, 9H, -$(CH_3)_3$), 2.45 (dd, J_1 = 16.8 Hz, J_2 = 7.2 Hz, 1H, -CH_2), 2.56 (dd, J_1 = 16.4 Hz, J_2 = 5.2 Hz, 1H, -CH_2), 3.51 (dd, J_1 = 6.8 Hz, J_2 = 5.2 Hz, 1H, -CH), 3.64 (dd, J_1 = 6.8 Hz, J_2 = 5.2 Hz, 2H, -CH_2), 4.34 (t, *J* = 5.6 Hz, 2H, -CH_2), 8.12 (s, 1H, adenine H), 8.21 (s, 1H, adenine H). ^{13}C NMR (101 MHz, CD_3OD, 298K) δ_C = 29.6, 41.5, 42.2, 45.6, 54.1, 83.5, 121.3, 144.3, 152.1, 154.9, 158.5, 173.4, 177.7. HRESI-MS calc. for $C_{15}H_{24}N_7O_3^+$ 350.1935, found 350.1929.

S0 (670 mg, 1.00 mmol), 2-(*1H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU, 379 mg, 1.00 mmol), **S4** (350 mg, 1.00 mmol) and triethylamine (0.30 mL, 2.15 mmol) were

S5

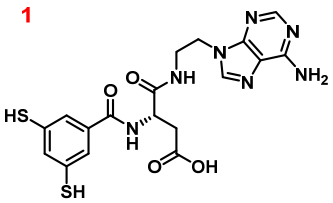


dissolved in 10 mL dry DMF and the mixture was stirred at room temperature overnight under nitrogen atmosphere. Solvent was evaporated under vacuum and the crude mixture was purified by flash column chromatography (SiO_2, 0-10% methanol in DCM), followed by removal of solvent to afford **S5** as a white powder.

Yield = 67%, 670 mg.

^{1}H NMR: (400 MHz, DMSO-d6, 298K) δ_H = 1.30 (s, 9H, -$(CH_3)_3$), 2.43 (dd, J_1 = 15.6 Hz, J_2 = 8.8 Hz, 1H, -CH_2), 2.57 (dd, J_1 = 16 Hz, J_2 = 5.6 Hz, 1H, -CH_2), 3.35-3.43 (m, 1H, -CH_2), 3.51-3.56 (m, 1H, -CH_2), 4.20-4.24 (m, 2H, -CH_2), 4.53 (ddd, J_1 = 13.6 Hz, J_2 = 8.4 Hz, J_3 = 4.8 Hz, 1H, -CH), 6.62 (s, 1H, ArH), 7.17-7.23 (m, 32H, ArH), 8.07 (t, J = 6.0 Hz, 1H, NH), 8.13 (d, J = 8.0 Hz, 1H, NH), 8.21 (s, 1H, adenine H), 8.27 (s, 1H, adenine H), 8.85 (br, 2H, -NH_2). ^{13}C NMR (101 MHz, DMSO-d6, 298K) δ_C = 30.8, 40.3, 46.2, 48.9, 53.2, 74.0, 83.2, 121.5, 130.0, 131.0, 132.4, 135.8, 136.5, 136.8, 144.0, 146.0, 146.6, 151.0, 152.2, 155.5, 168.2, 172.4, 173.6. HRESI-MS calc. for $C_{60}H_{56}N_7O_4S_2^+$ 1002.3830, found 1002.3848.

1



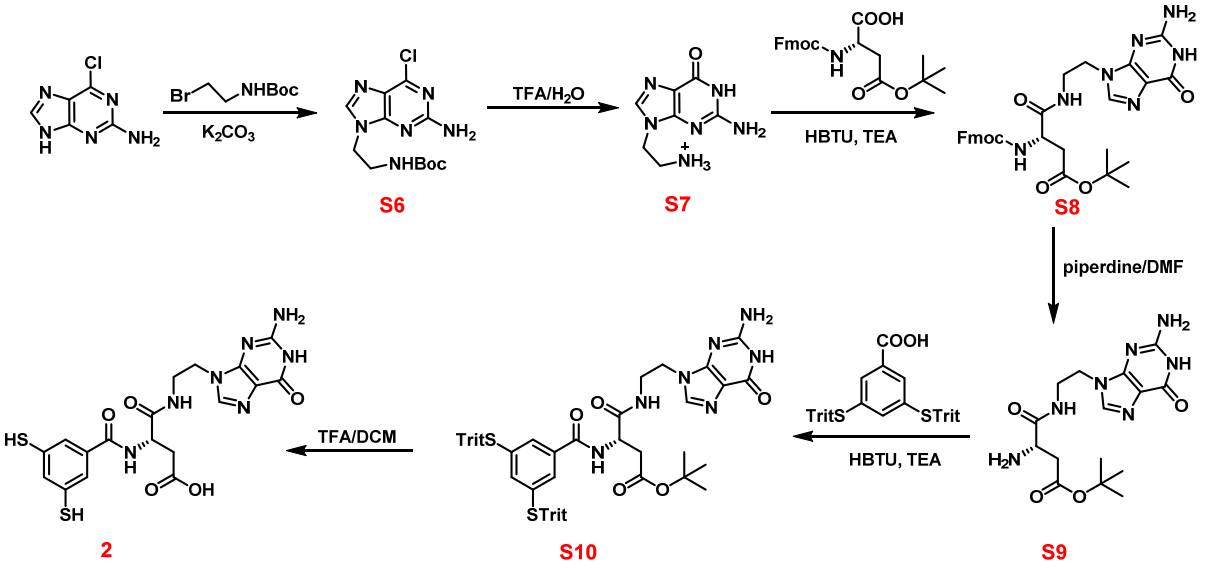
S5 (100 mg, 0.10 mmol) was dissolved in 10 mL 50% TFA in DCM and stirred for 12 h at room temperature under N_2. Et_3SiH (0.50 mL, 3.1 mmol) was added to the reaction mixture and further stirred for another hour. Solvents were evaporated under vacuum and the crude mixture was washed with hexane (10 mL x 2). The product was purified by reverse phase flash column chromatograph (RP C18, 0-90% acetonitrile in water with 0.1% TFA), and the desired product **1** was obtained after lyophilization as a white powder.

Yield = 35%, 16 mg.

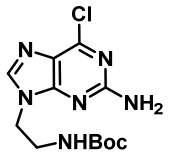
^{1}H NMR: (400 MHz, CD_3OD, 298K) δ_H = 2.80 (dd, J_1 = 16.8 Hz, J_2 = 8.0 Hz, 1H, -CH_2), 2.91 (dd, J_1 = 16.8 Hz, J_2 = 8.0 Hz, 1H, -CH_2), 3.75-3.78 (m, 2H, -CH_2), 4.49 (dd, J_1 = 9.6 Hz, J_2 = 4.8 Hz, 2H, -CH_2), 4.78-4.81 (m, 1H, -CH), 7.46-7.47 (m, 1H, ArH), 7.53-7.54 (m, 2H, ArH), 8.39 (s, 1H, adenine H), 8.40 (s, 1H, adenine H). ^{13}C NMR (101 MHz, CD_3OD, 298K) δ_C = 36.4, 40.2, 45.0, 52.0, 119.7, 125.2, 132.0, 135.3, 136.4, 145.5, 145.8, 150.7, 151.9, 168.6, 173.6, 173.9. HRESI-MS calc. for $C_{18}H_{18}N_7O_4S_2^-$ 460.0862, found 460.0850.

Chapter 2

Synthesis of building block **2**



S6

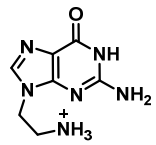


The procedure for the synthesis of **S6** was the same as that described for the synthesis of **S1.**

Yield = 35%, 1.64 g.

^{1}H NMR: (400 MHz, DMSO-d6, 298K) δ_H = 1.29 (s, 9H, Boc), 3.29-3,33 (m, 2H, -CH_2), 4.06 (t, *J* = 5.9 Hz, 2H, -CH_2), 6.84 (s, 2H, -NH_2), 6.93 (t, *J* = 5.2 Hz, 1H, -CONH), 7.95 (s, 1H, guanine H). ^{13}C NMR (101 MHz, DMSO-d6, 298K) δ_C = 21.6, 22.2, 28.1, 43.1, 77.9, 123.5, 143.3, 149.1, 154.3, 155.5, 159.7. HRESI-MS calc. for $C_{12}H_{18}ClN_6O_2^+$ 313.1174, found 313.1177.

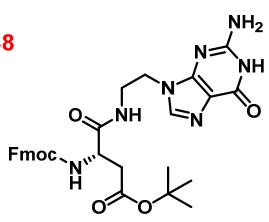
S7



S6 (0.50 g, 1.6 mmol) was dissolved in 10 mL 50% TFA/water under a nitrogen atmosphere and the mixture was stirred for 24 h. After that, the solvents were evaporated under reduced pressure and the crude product was precipitated by addition of diethylether. The white powder was collected by filtration. Yield = 98%, 0.46 g.

^{1}H NMR: (400 MHz, DMSO-d6, 298K) δ_H = 3.32 (q, *J* = 5.5 Hz, 2H, -CH_2), 4.32 (t, *J* = 5.7 Hz, 2H, -CH_2), 7.09 (s, 2H, -NH_2), 8.28 (s, 3H, -NH_3^+), 8.42 (s, 1H, guanine H), 11.46 (s, 1H, guanine NH). ^{13}C NMR (101 MHz, DMSO-d6, 298K) δ_C = 37.8, 42.0, 112.1, 137.7, 150.6, 154.8, 155.2. HRESI-MS calc. for $C_7H_{11}N_6O^+$ 195.0989, found 195.0989.

S8

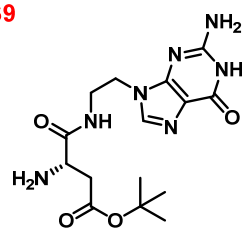


The procedure for the synthesis of **S8** was the same as that described for the synthesis of **S3**.

Yield = 43%, 0.25 g.

^{1}H NMR: (400 MHz, DMSO-d6, 298K) δH = 1.31 (s, 9H, -$(CH_3)_3$), 2.41 (dd, J_1 = 16.0 Hz, J_2 = 9.5 Hz, 1H, -CH_2), 2.59 (dd, J_1 = 16.0 Hz, J_2 = 4.9 Hz, 1H, -CH_2), 3.28-3.49 (m, 2H, -CH_2), 3.98 (t, J = 5.7 Hz, 2H, -CH_2), 4.16-4.36 (m, 4H, -CH&CH_2), 6.50 (s, 2H, -NH_2), 7.29 (t, J = 7.5 Hz, 2H, ArH), 7.39 (t, J = 7.5 Hz, 2H, ArH), 7.57 (s, 1H, guanine H), 7.61 (d, J = 8.3 Hz, 1H, -CONH), 7.69 (t, J = 6.4 Hz, 2H, ArH), 7.86 (d, J = 7.6 Hz, 2H, ArH), 8.11 (s, 1H, -CONH), 10.62 (s, 1H, guanine NH). ^{13}C NMR (101 MHz, DMSO-d6, 298K) δ_C = 27.7, 37.5, 42.2, 45.6, 46.6, 51.6, 65.8, 80.2, 116.5, 120.1, 125.3, 127.1, 127.7, 137.7, 140.7, 143.7, 143.9, 151.2, 153.6, 155.8, 156.8, 169.4, 170.9. HRESI-MS calc. for $C_{30}H_{34}N_7O_6^+$ 588.2565, found 588.2565.

S9



The procedure for the synthesis of **S9** was the same as that described for the synthesis of **S4**.

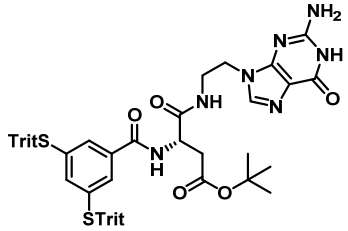
Yield = 67%, 0.24 g.

^{1}H NMR: (400 MHz, DMSO-d6, 298K) δ_H = 1.39 (s, 9H, -$(CH_3)_3$), 2.68 (qd, J_1 = 17.6 Hz, J_2 = 6.2 Hz, 2H, -CH_2), 3.42 (dd, J_1 = 13.9 Hz, J_2 = 5.5 Hz, 1H, -CH_2), 3.51 (dq, J_1 = 11.7 Hz, J_2 = 5.9 Hz, 1H, -CH_2), 3.95 (dd, J_1 = 8.1 Hz, J_2 = 4.5 Hz, 1H, -CH), 4.02 (t, J = 5.6 Hz, 2H, -CH_2), 6.59 (s, 2H, -NH_2), 7.77 (s, 1H, guanine H), 8.19 (s, 2H, NH_2), 8.57 (t, J = 5.8 Hz, 1H, -CONH), 10.80 (s, 1H, guanine NH). ^{13}C NMR (101 MHz, DMSO-d6, 298K) δ_C = 27.7, 36.1, 42.4, 48.8, 64.9, 81.6, 115.6, 137.6, 151.2, 15.9, 156.5, 167.9, 168.6. HRESI-MS calc. for $C_{15}H_{24}N_7O_4^+$ 366.1884, found 366.1890.

The procedure for the synthesis of **S10** was the same as that described for the synthesis of **S5.**

Yield = 54%, 0.55 g.

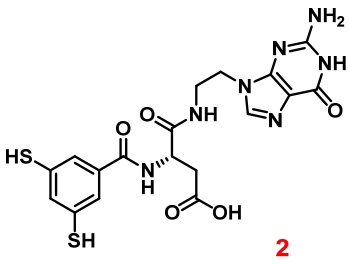
S10



^{1}H NMR: (400 MHz, DMSO-d6, 298K) δ_H = 1.30 (s, 9H, -$(CH_3)_3$), 2.46-2.51 (m, 1H, -CH_2), 2.61 (dd, J_1 = 16.0 Hz, J_2 = 4.8 Hz, 1H, -CH_2), 3.32-3.40 (m, 2H, -CH_2), 3.96 (t, J = 5.8 Hz, 2H, -CH_2), 4.61-4.55 (m, 1H, -CH), 6.44 (s, 2H, -NH_2), 6.60 (s, 1H, ArH), 7.14-7.24 (m, 32H, ArH), 7.57 (s, 1H, guanine H), 8.07 (t, J = 5.8 Hz, 1H, CONH), 8.16 (d, J = 8.1 Hz, 1H, CONH), 10.55 (s, 1H, guanine NH). ^{13}C NMR (101 MHz, DMSO-d6, 298K) δ_C = 27.7, 37.3, 42.2, 50.2, 70.9, 80.1, 126.9, 127.8, 129.3, 132.7, 133.4, 133.7, 137.6, 140.9, 143.5, 151.1, 153.5, 156.7, 164.1, 169.3, 170.5. HRESI-MS calc. for $C_{60}H_{56}N_7O_5S_2^+$ 1018.3779, found 1018.3844.

The procedure for the synthesis of **2** was the same as that described for the synthesis of **1**.

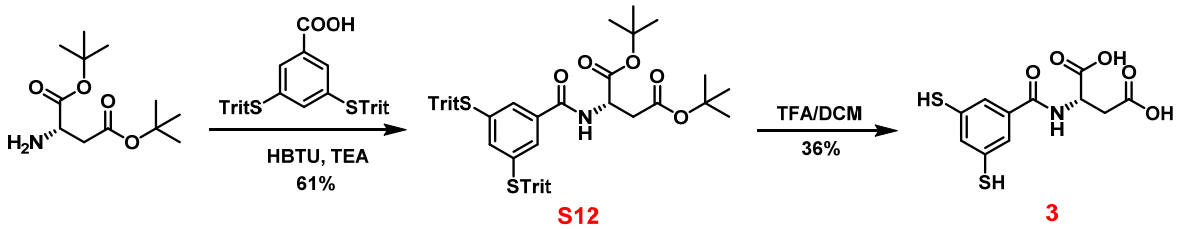
Yield = 34%, 16 mg.



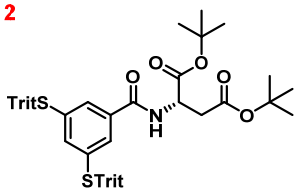
2

^{1}H NMR: (400 MHz, CD_3OD, 298K) δ_H = 2.77 (dd, J_1 = 16.8 Hz, J_2 = 7.8 Hz, 1H, -CH_2), 2.89 (dd, J_1 = 16.8 Hz, J_2 = 5.9 Hz, 1H, -CH_2), 3.64 (t, J = 5.5 Hz, 2H, -CH_2), 4.28 (t, J = 5.5 Hz, 2H, -CH_2), 4.74 (dd, J_1 = 7.8, J_2 = 5.9 Hz, 1H, -CH), 7.35 (t, J = 1.7 Hz, 1H, ArH), 7.44 (d, J = 1.7 Hz, 2H, ArH), 8.46 (s, 1H, guanine H). No ^{13}C NMR could be obtained due to the poor solubility of **2**. HRESI-MS calc. for $C_{18}H_{20}N_7O_5S_2^+$ 478.0962, found 478.0959.

Synthesis of building block 3



S12

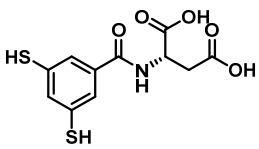


S0 (670 mg, 1.00 mmol), 2-(*1H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU, 379 mg, 1.00 mmol), *L*-aspartic acid di-*tert*-butyl ester hydrochloride (281 mg, 1.00 mmol) and triethylamine (0.30 mL, 2.15 mmol) were dissolved in 10 mL dry DMF and the mixture was stirred at room temperature overnight under N_2. The solvent was evaporated under vacuum and the crude mixture was purified by flash column chromatography (SiO_2, 0-10% methanol in DCM), followed by removal of solvent to afford derivative **S12** as a white powder.

Yield = 61%, 547 mg.

^{1}H NMR: (400 MHz, $CDCl_3$, 298K) δ_H = 1.45 (s, 9H, -$(CH_3)_3$), 1.47 (s, 9H, -$(CH_3)_3$), 2.67 (dd, J_1 = 16.8 Hz, J_2 = 4.4 Hz, 1H, -CH_2), 2.83 (dd, J_1 = 17.2 Hz, J_2 = 4.8 Hz, 1H, -CH_2), 4.65-4.69 (m, 1H, -CH), 6.33 (d, J = 8.0 Hz, 1H, -CONH), 6.92 (d, J = 1.6 Hz, 2H, ArH), 7.06 (t, J = 1.7 Hz, 1H, ArH), 7.14-7.22 (m, 18H, ArH), 7.28-7.35 (m, 12H, ArH). ^{13}C NMR (101 MHz, $CDCl_3$, 298K) δ_C = 29.6, 30.6, 30.8, 40.2, 52.1, 74.1, 83.9, 84.8, 129.5, 130.4, 130.6, 132.5, 135.4, 136.3, 137.4, 145.5, 146.7, 168.2, 172.1, 172.8. HRESI-MS calc. for $C_{57}H_{56}NO_5S_2^+$ 898.3594, found 898.3587.

3



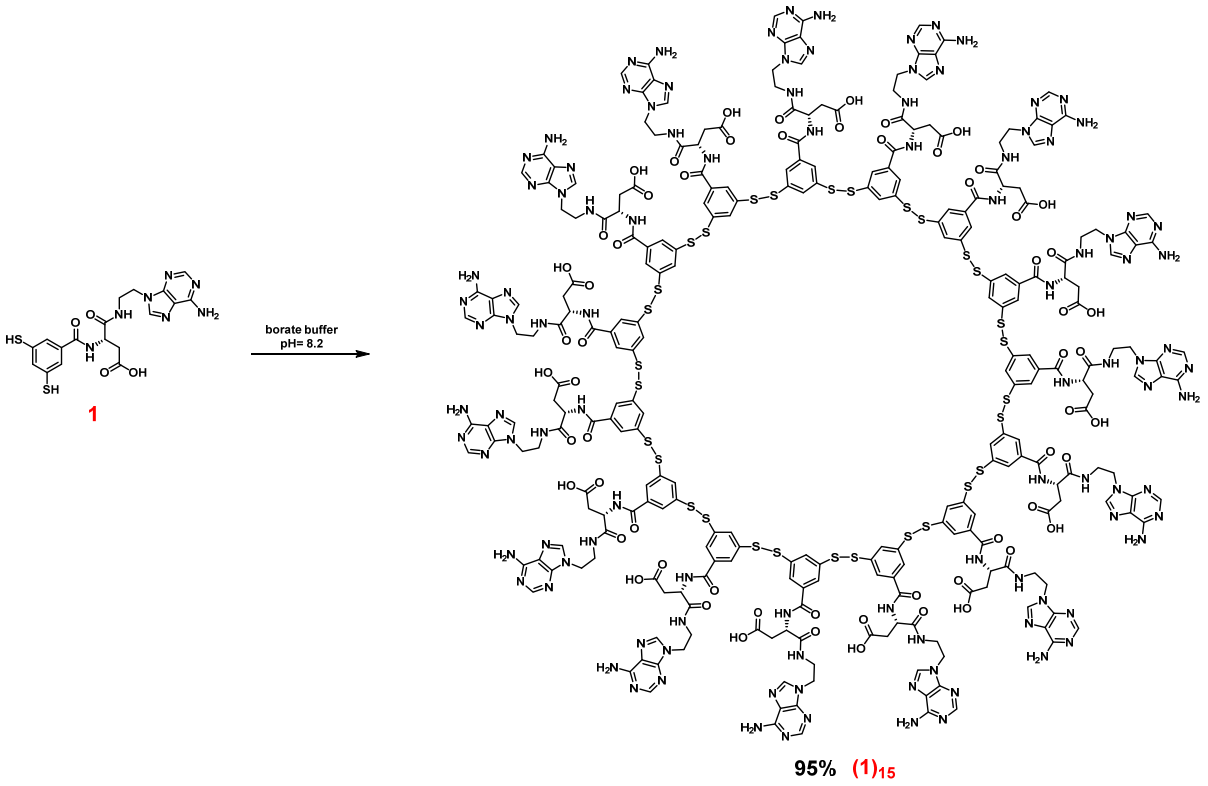
The procedure for the synthesis of **3** was the same as that described for the synthesis of **1**.

Yield = 36%, 12 mg.

^{1}H NMR: (400 MHz, CD_3OD, 298K) δ_H = 2.88 (dd, J_1 = 16.7 Hz, J_2 = 7.5 Hz,

1H, -CH_2), 2.98 (dd, J_1 = 16.7 Hz, J_2 = 5.3 Hz, 1H, -CH_2), 4.92 (dt, J_1 = 7.5 Hz, J_2 = 5.1 Hz, 1H, -CH), 7.37 (t, J = 1.8 Hz, 1H, ArH), 7.46 (d, J = 1.7 Hz, 2H, ArH). ^{13}C NMR (101 MHz, CD_3OD, 298K) δ_C = 37.9, 52.2, 126.4, 133.2, 136.6, 137.9, 175.2, 175.3. HRESI-MS calc. for $C_{11}H_{12}NO_5S_2^+$ 302.0151, found 302.0147.

Synthesis of foldamer $(1)_{15}$



Building block **1** (50 mg) was dissolved at a 2.0 mM concentration in borate buffer (50 mM, pH=8.2) containing 1.0 M NaCl. The reaction mixture was stirred at room temperature under air and the library was monitored by UPLC. After all of monomer was consumed, the product was purified by reverse phase flash column chromatograph (RP C18, 0-90% acetonitrile in water with 0.1% TFA), and the desired foldamer $(\mathbf{1})_{15}$ was obtained after lyophilization as a white powder.

Yield = 94%, 47 mg.

^{1}H NMR: (500 MHz, D_2O, 298K) δ_H = 2.35 (d, J = 13.2 Hz, 1H, -CH_2), 2.64 (d, J = 17.2 Hz, 1H, -CH_2), 2.74 (d, J = 16.3 Hz, 2H, -CH_2), 2.81-2.94 (m, 4H, -CH_2), 3.04-3.21 (m, 2H, -CH_2), 3.46 (d, J = 11.5 Hz, 1H, -CH_2), 3.85 (dd, J_1 = 33.0, J_2 = 13.8 Hz, 2H, -CH_2), 3.93-3.98 (m, 3H, -CH_2), 4.13 (s, 1H, -CH), 4.52 (s, 2H, -CH_2), 4.82 (s, 1H, -CH), 5.06 (s, 1H, -CH), 5.72 (s, 1H, ArH), 6.37 (s, 1H, ArH), 6.85 (s, 2H, ArH), 7.05 (s, 1H, ArH), 7.11 (s, 1H, ArH), 7.12 (s, 2H, ArH & adenien H), 7.36 (s, 1H, ArH), 7.45 (s, 1H, ArH), 8.07 (s, 1H, adenine H), 8.13 (s, 1H, adenine H), 8.31 (s, 1H, adenine H), 8.33 (s, 1H, adenine H), 8.42 (s, 1H, adenine H), 8.92 (s, 2H, -CONH), 9.04 (d, J = 7.6 Hz, 1H, -CONH).

2.5 Appendix

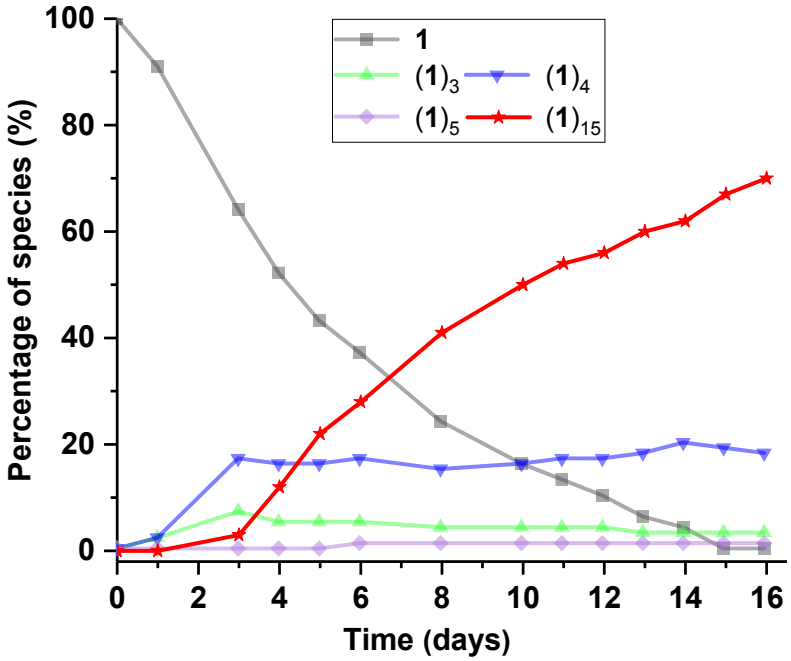


Figure S2.1. Kinetic profile of the library made from 1.0 mM building block **1** in borate buffer (pH = 8.2, 50 mM) under continuous stirring.

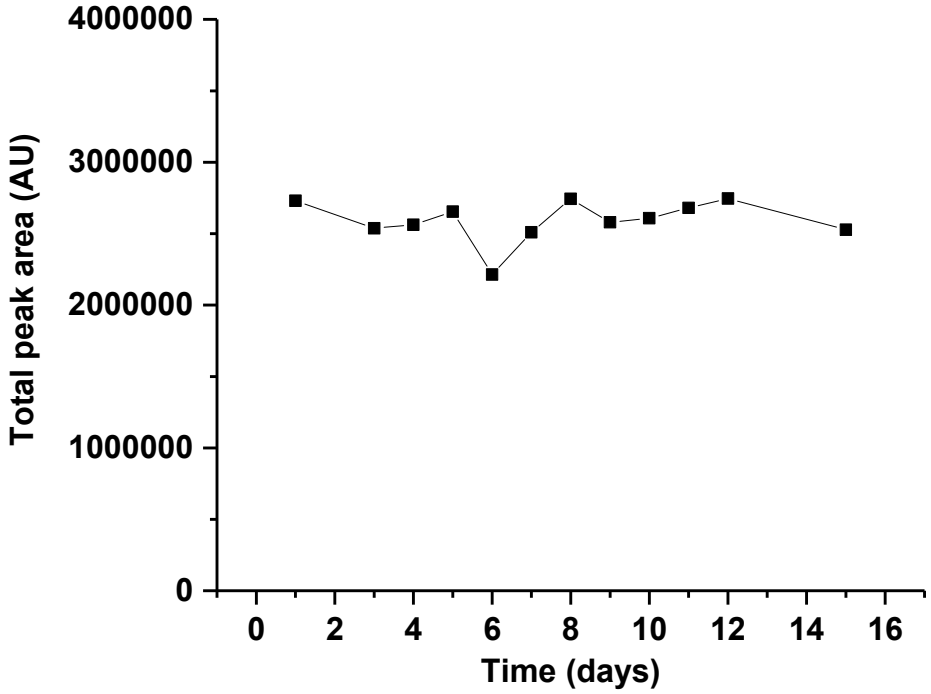


Figure S2.2. Total UPLC peak area of library made from building block **1** monitoring the library shown in Fig. S2.1.

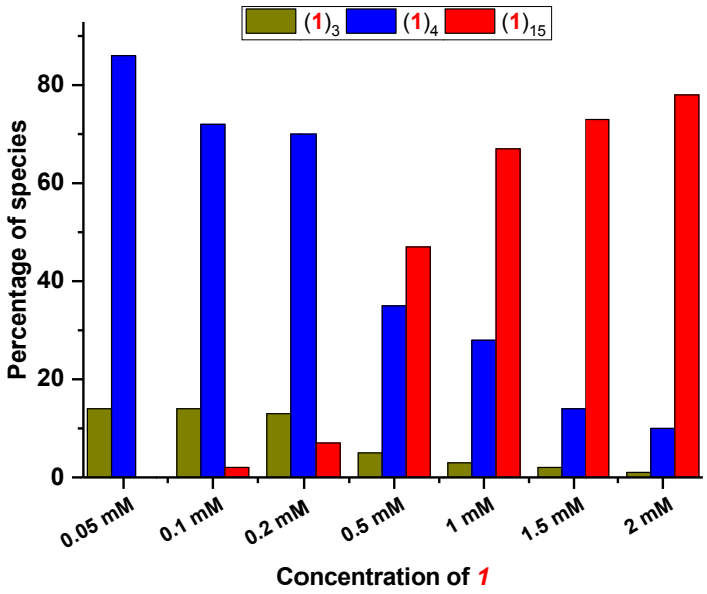


Figure S2.3. Effect of the concentration of building block **1** on the product distribution. Libraries were prepared by dissolving different amounts of building block **1** in borate buffer (pH = 8.2, 50 mM). The distribution of the products was monitored by UPLC after the libraries had reached equilibrium.

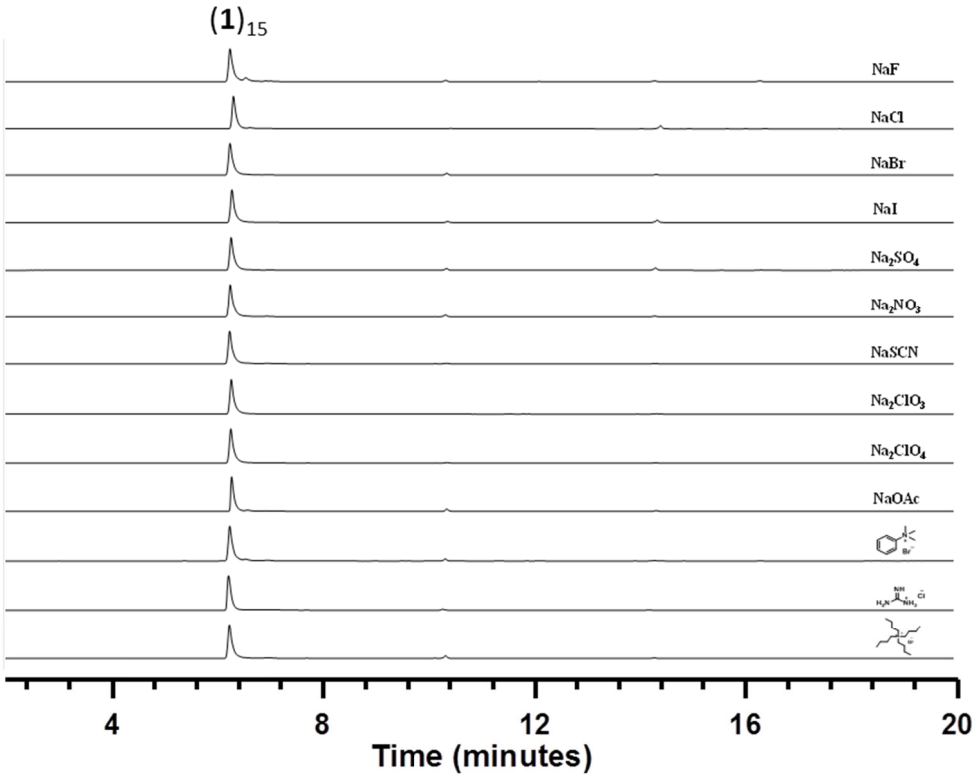


Figure S2.4. Effect of salts on the formation of foldamer $(\mathbf{1})_{15}$. The libraries were prepared by dissolving building block **1** (0.50 mM) in borate buffer containing 1.0 M of a specific salt. After the libraries had reached equilibrium the distribution of the products was monitored by UPLC.

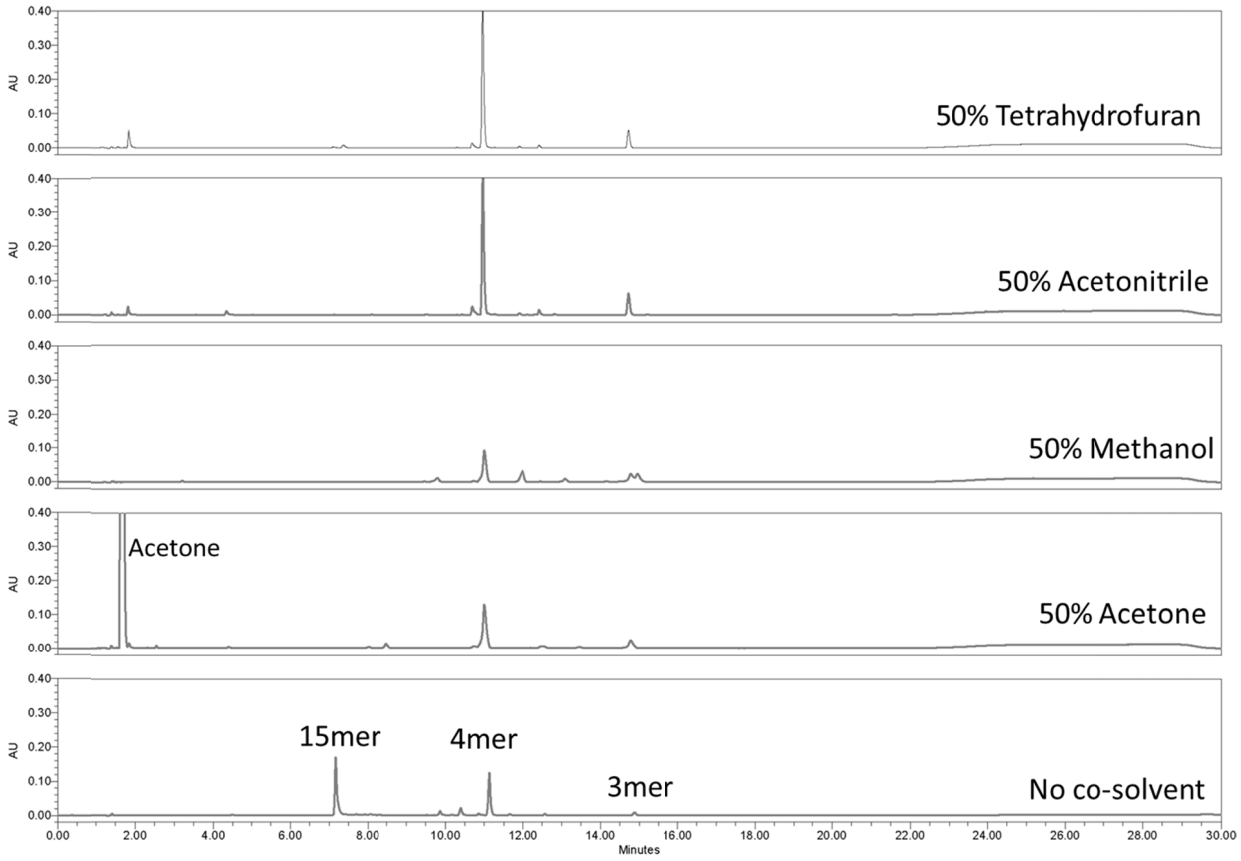


Figure S2.5. Effect of co-solvents on prodcut distribution of the library made from building block **1**. The libraries were prepared by dissolving building block **1** (0.50 mM) in borate buffer with 50% (volume) organic co-solvent. UPLC spectra from bottom to the top: No co-solvents, 50% acetone, 50% methanol, 50% acetonitrile and 50% tetrahydrofuran.

Proton NMR assignment of foldamer $(1)_{15}$

The assignment of the proton NMR was based on 2D NMR and the crystal structure of the foldamer.

Proton 1'' is the only proton which is in close contact with two protons belonging to other aryl rings (3 and 1'), in the crystal structure and the NOESY NMR shows the corresponding cross-peaks.

Assignments of adenine signals to A8/A9 protons were based on the ^{1}H, ^{13}C HMBC experiment:

In the structure of adenine, only A8 protons give rise to cross-peaks due to correlations with C-A5 (least shifted on the map) in HMBC spectrum (blue spots). Protons A9 and A8 correlate (pairwise) with the same carbon (C-A4) in the HMBC spectrum (about 150 ppm), and the A9 protons correlate with the C-A6 (most shifted carbons). Taken together, these cross-peaks allowed us to unequivocally assign the proton signals of the adenines.

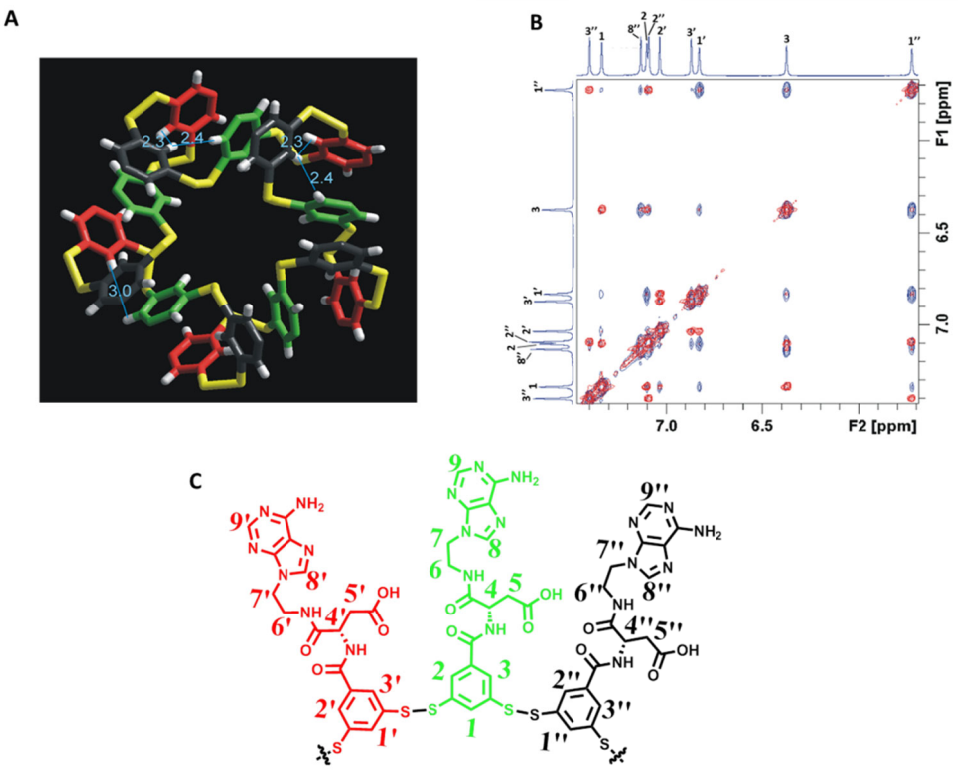


Figure S2.6. (A) Core part of the foldamer indicating the distance between protons as observed in the crystal structure. (B) Superimposed COSY (red) and NOESY (blue) spectra. (C) Repeating unit of $(\mathbf{1})_{15}$ colored as in A.

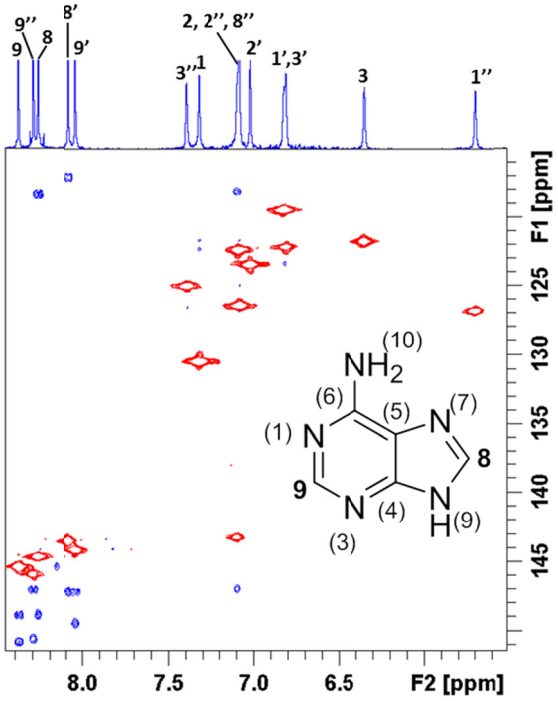


Figure S2.7. Part of the HMBC map (blue) and the superimposed HSQC map (red) of $(\mathbf{1})_{15}$ showing the assignment of protons 8 and 9.

Distinction between 2 (or 3) and 1 was made on the basis of ^{1}H, ^{13}C-HMBC: Only H2's and H3's give rise to cross-peaks in HMBC with amide C=O carbons. That allowed the assignments of proton signals to either position 1 or 2/3 in each of the aryl rings.

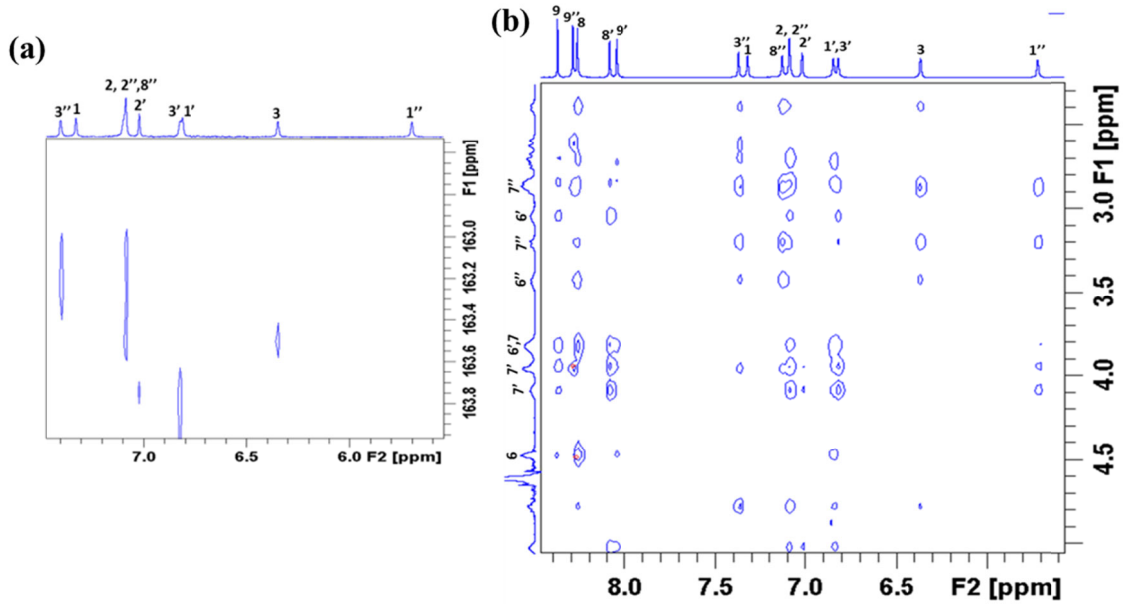


Figure S2.8. (a) Part of the HMBC map of $(\mathbf{1})_{15}$ showing the contact between protons 2, 3 and amide C=O, (b) Parts of the NOESY maps of of $(\mathbf{1})_{15}$ showing the contact between protons (6 and 7) and aromatic protons.

Assignments of ethylene-bridge protons (6 and 7) were made on the basis of discrete NOE correlations with aryl protons at various conformations available in the structural model. Assignments of these protons to a specific carbon of ethylene was made on the basis of ^{1}H,^{13}C HSQC spectra (diastereotopic methylene proton signals, except 6 and 7, correlate with the same carbon). In most cases, the correlations observed in the NOESY map for 6 and 7 protons were with aryl protons of a different monomer unit (different color). Strong correlations of ethylene protons were also observed with adenine 8 protons with the same monomer unit.

The Asn 4 and 5 protons (diastereotopic, strongly overlapping with each other) were assigned by various contacts with either ethylene-linker protons 6, 7 or adenine proton 8. One NH of Asn appeared to be most deeply buried inside the foldamer and, thus, exchanged only slowly with D_2O. Indeed, its signal correlates with that of proton 4 in the COSY spectrum. The other NH's are either exchanged completely, or do not correlate with proton 4, neither in COSY nor NOESY experiments, even when they are recorded in H_2O solution. Although the α-β-β'-triads were assigned for each subunit, attempts to do it on the basis of H-C correlations failed, despite several approaches, due to lack of any correlations between α-protons with amide C=O in the HMBC spectrum.

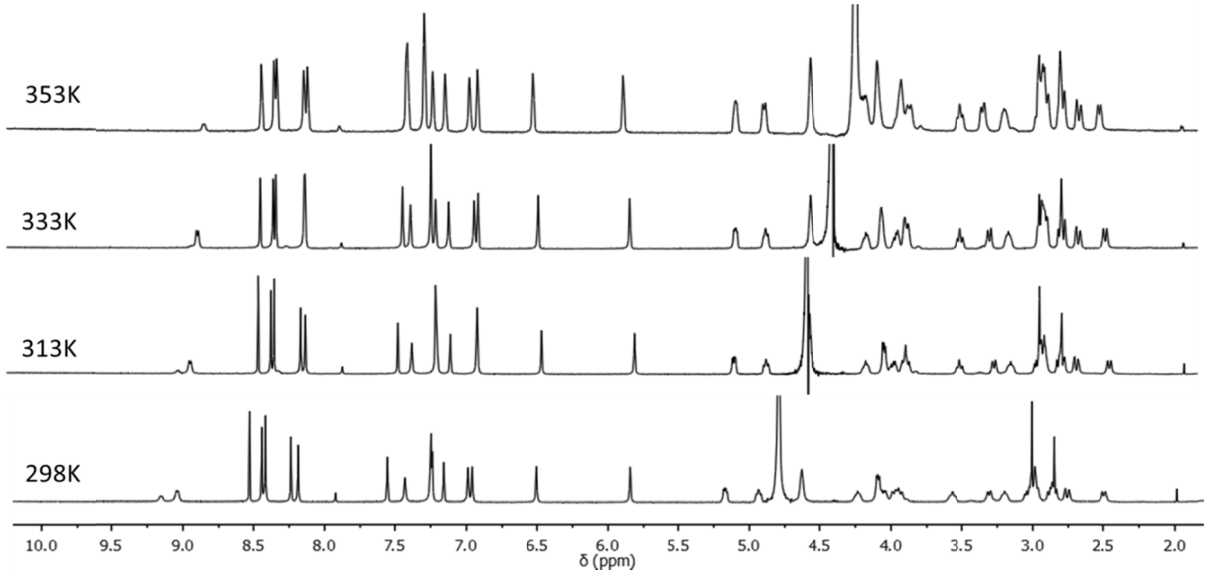


Figure S2.9. Variable temperature ^{1}H-NMR spectra of foldamer $(\mathbf{1})_{15}$ in D_2O at 353 K, 333 K, 313 K and 298 K (600 MHz).

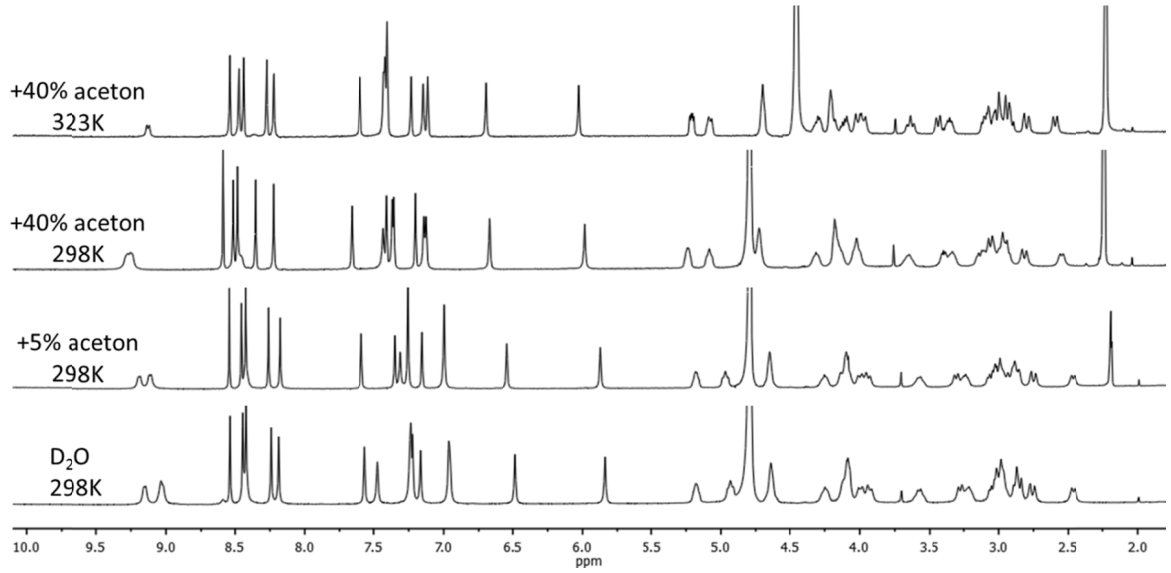


Figure S2.10. Variable temperature ^{1}H-NMR spectra of the foldamer $(\mathbf{1})_{15}$ in D_2O at 323 K with 40% acetone; 298 K with 40% acetone; 298 K with 5% acetone and 298 K in D_2O (500 MHz).

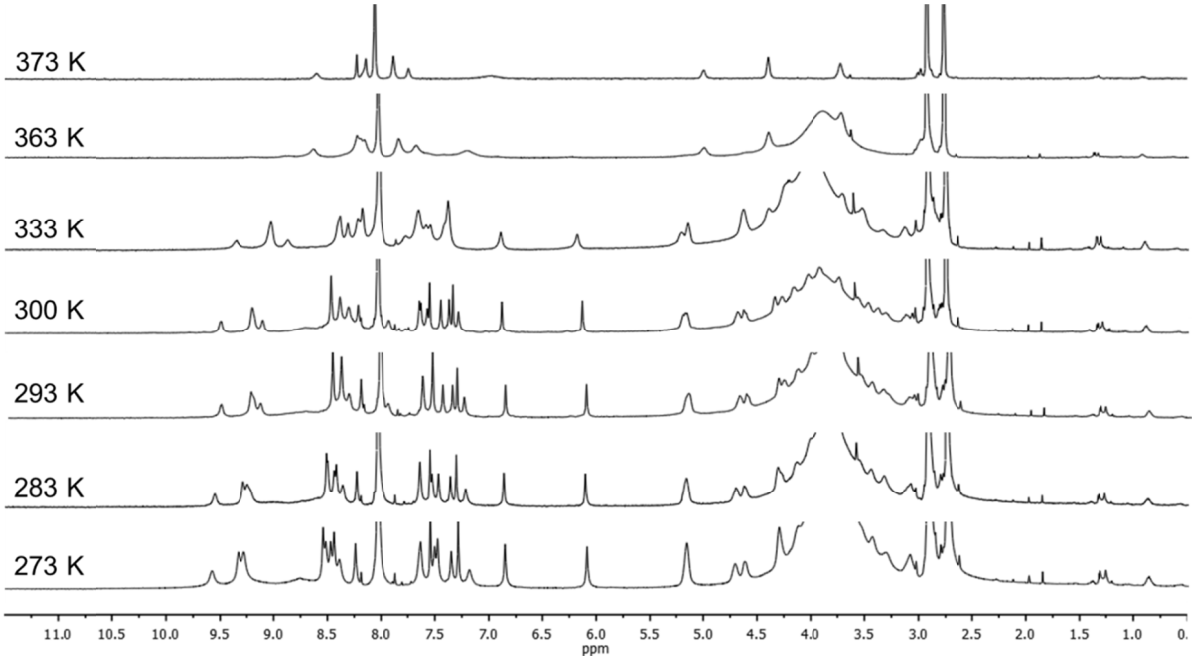


Figure S2.11. Variable temperature ^{1}H-NMR spectra of the foldamer $(\mathbf{1})_{15}$ in DMF-d_7 at 373 K, 363 K, 333 K, 300 K, 293 K, 283 K and 273 K (600 MHz).

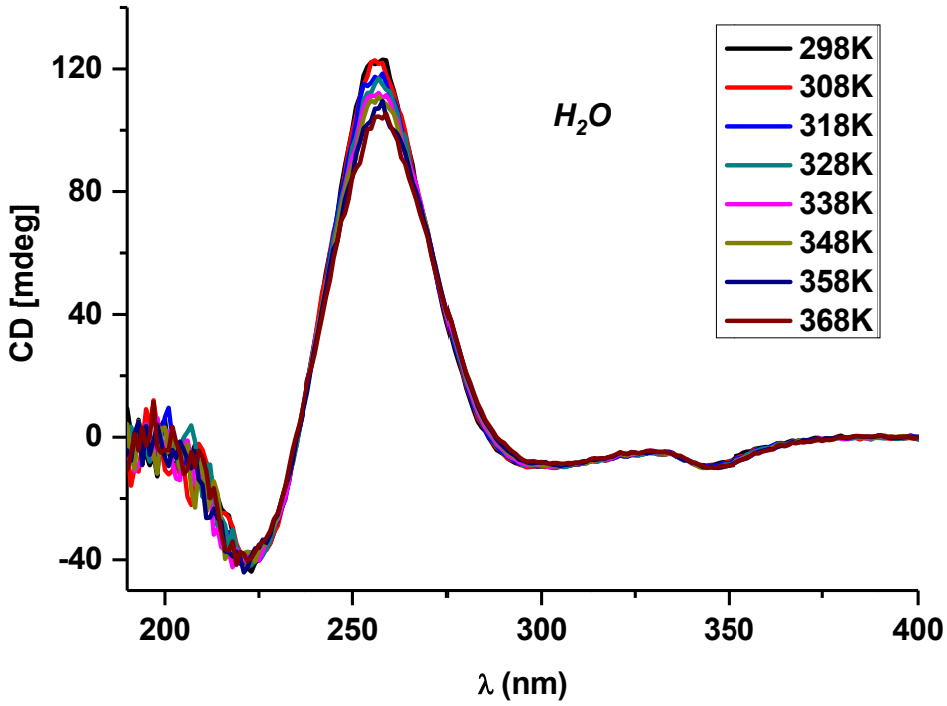


Figure S2.12. CD spectra of foldamer $(\mathbf{1})_{15}$ in water at different temperatures.

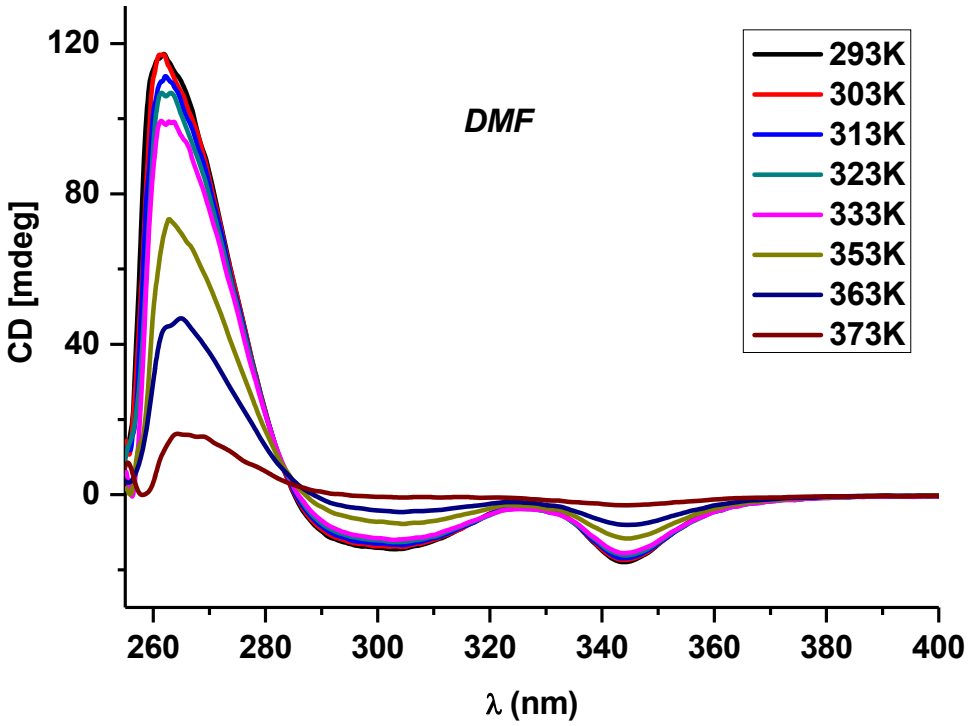


Figure S2.13. CD spectra of foldamer $(\mathbf{1})_{15}$ in DMF at different temperatures.

Details of single crystal structure of $(\mathbf{1})_{15}$

Single crystals were obtained by slow diffusion of acetone into a water solution of foldamer $(\mathbf{1})_{15}$. Data collection of foldamer was performed at the X-ray diffraction beamline (XRD1) of the Elettra Synchrotron of Trieste (Italy), with a Pilatus 2M image plate detector. Complete dataset was collected at 100 K (nitrogen stream supplied by an Oxford Cryostream 700) with a monochromatic wavelength of 0.700 Å, respectively, through the rotating crystal method. The crystal was dipped in N-paratone and mounted on the goniometer head with a nylon loop. The diffraction data were indexed, integrated and scaled using XDS[38]. The structure was solved by direct methods using SIR2014[39] and subsequent Fourier analysis and refinements with the full-matrix least-squares method based on F^2 were performed with SHELXL[40]. Anisotropic thermal motion was allowed for all non-H atoms, except for the oxygens of lattice water molecules. Hydrogen atoms were placed at calculated positions and no H atoms were assigned to water molecules. The unit cell presents almost 39.7% of void, likely occupied by highly disordered solvent molecules, and the program Squeeze was applied to the data set to take into account this fact. Graphics were drawn with program Diamond[41].

Crystal data: $C_{270}H_{240}N_{105}O_{60}S_{30}.27.5(H_2O)$, M = 7381.46, monoclinic, space group $P\ 2_12_12$, a = 100.02(2), b = 18.644(4), c = 29.230(6) Å, V = 54507(19) Å^3, Z = 4, D_c = 0.892 g/cm^3, μ(Mo-Kα) = 0.142

mm^{-1}, *F*(000) = 15108, θ range = 0.69 to 19.55°. Final *R*1 = 0.1336, *wR*2 = 0.3475, *S* = 1.253 for 4839 parameters and 89211 reflections, 47051 unique [R(int) = 0.0270], of which 22953 with $I > 2\sigma(I)$, max positive and negative peaks in Δ*F* map 0.449 and -0.457 e. $Å^{-3}$.

2.6 MS spectrum

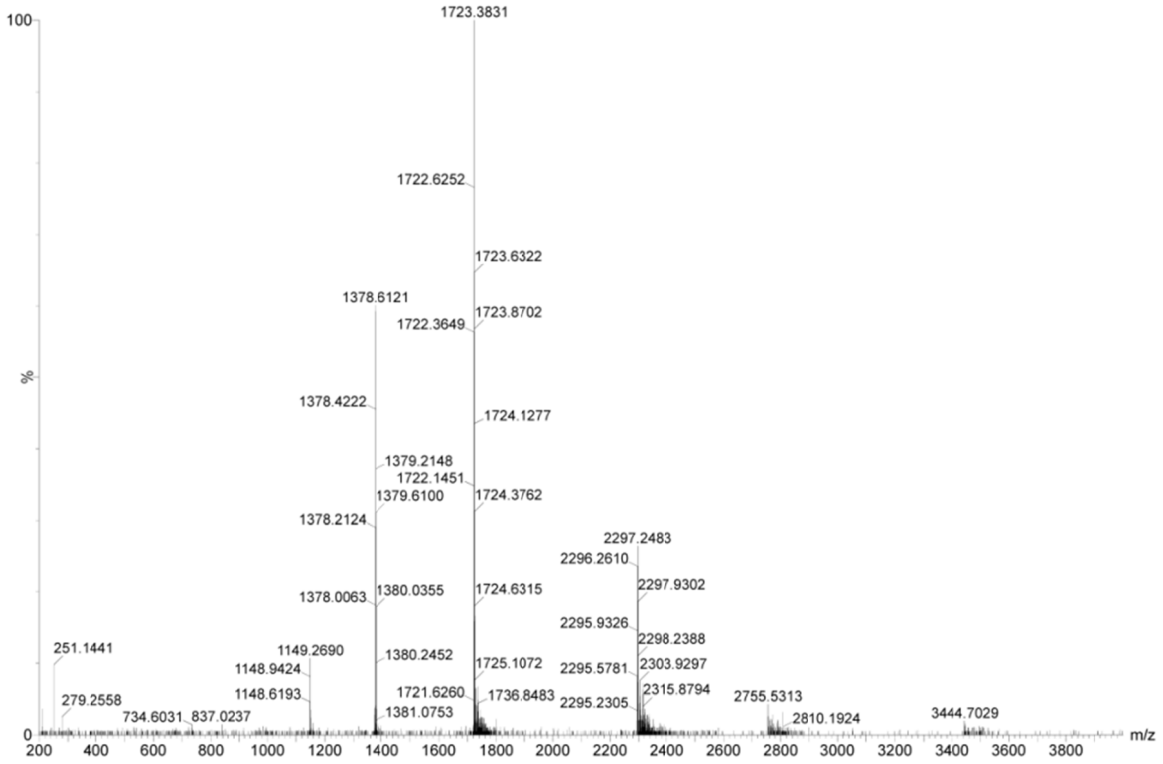


Figure S2.14. Mass spectrum of $(\mathbf{1})_{15}$ (retention time 5.6 min in Fig. 2.1) from the LC-MS analysis of a DCL made from **1**. Calculated m/z: 2296.40 $[M+3H]^{3+}$, 1722.55 $[M+4H]^{4+}$, 1378.24 $[M+5H]^{5+}$, 1148.70 $[M+6H]^{6+}$; observed m/z: 2297.25 $[M+3H]^{3+}$, 1723.38 $[M+4H]^{4+}$, 1378.61 $[M+5H]^{5+}$, 1149.27 $[M+6H]^{6+}$.

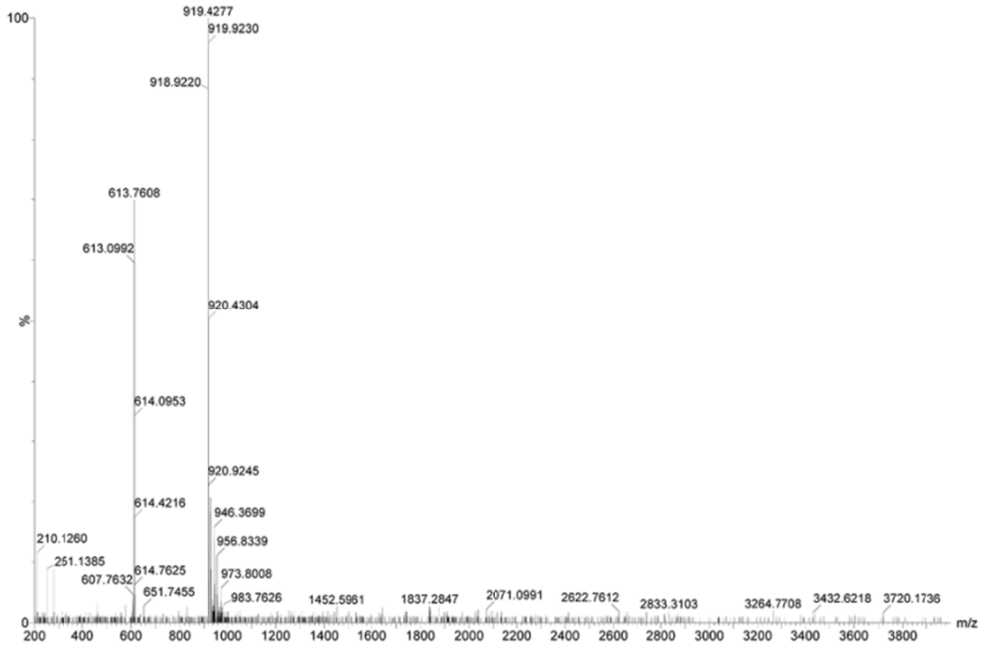


Figure S2.15. Mass spectrum of $(\mathbf{1})_4$ (retention time 7.7 min in Fig. 2.1) from the LC-MS analysis of a DCL made from **1**. Calculated m/z: 919.16 $[M+2H]^{2+}$, 613.11 $[M+3H]^{3+}$; observed m/z: 919.43 $[M+2H]^{2+}$, 613.76 $[M+3H]^{3+}$.

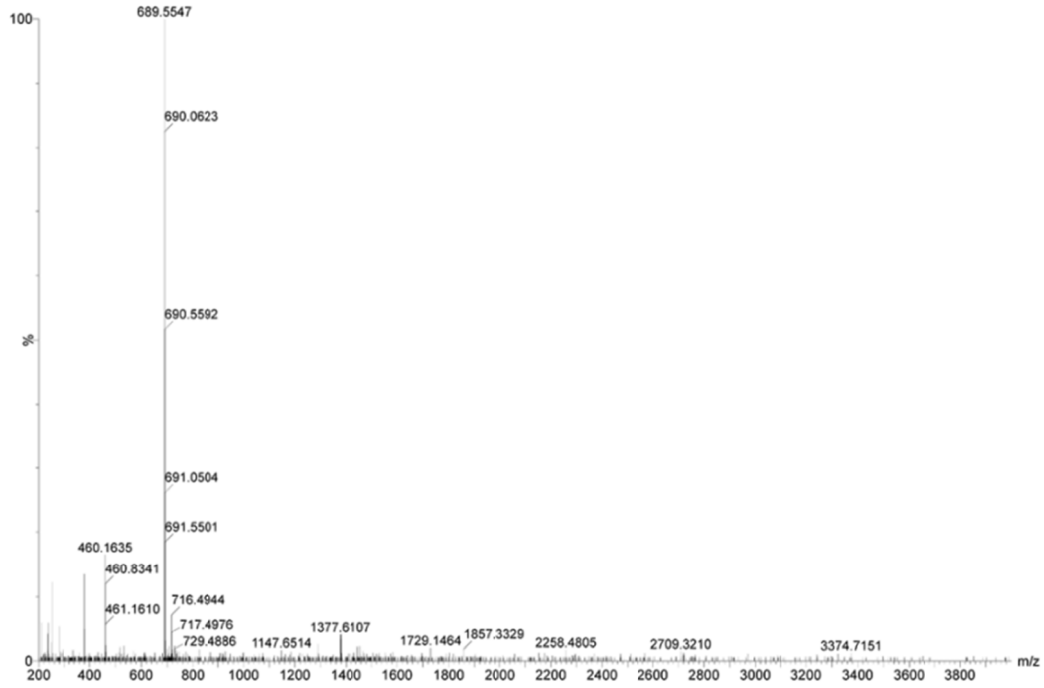


Figure S2.16. Mass spectrum of $(\mathbf{1})_3$ (retention time 9.6 min in Fig. 2.1) from the LC-MS analysis of a DCL made from **1**. Calculated m/z: 689.63 $[M+2H]^{2+}$, 460.09 $[M+3H]^{3+}$; observed m/z: 689.55 $[M+2H]^{2+}$, 460.16 $[M+3H]^{3+}$.

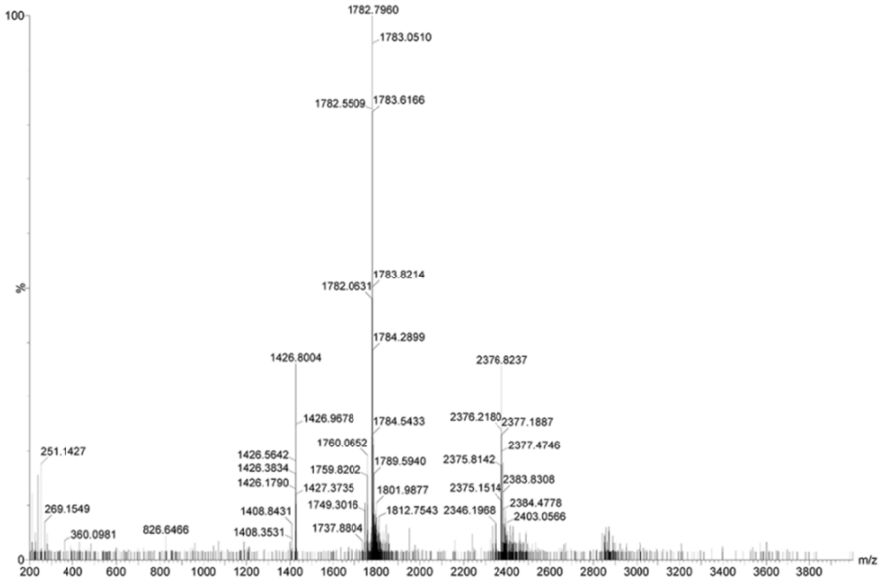


Figure S2.17. Mass spectrum of $(\mathbf{2})_{15}$ (retention time 6.1 min in Fig. 2.7c) from the L**C**-MS analysis of a DCL made from **2**. calculated m/z: 2376.37 $[M+3H]^{3+}$**,** 1782.53 $[M+4H]^{4+}$, 1426.23 $[M+5H]^{5+}$; observed m/z: 2376.82 $[M+3H]^{3+}$, 1782.80 $[M+4H]^{4+}$, 1426.80 $[M+5H]^{5+}$.

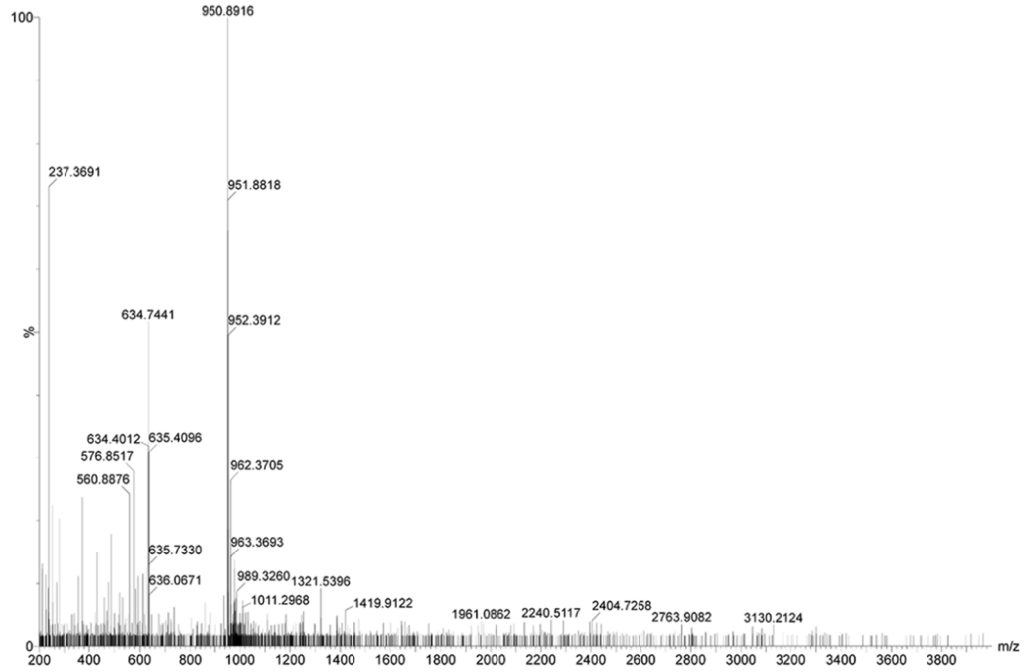


Figure S2.18. Mass spectrum of $(\mathbf{2})_4$ (retention time 11.8 min in Fig. 2.7c) from the L**C**-MS analysis of a DCL made from **2**. calculated m/z: 951.15 $[M+2H]^{2+}$, 634.44 $[M+3H]^{3+}$; observed m/z: 950.89 $[M+2H]^{2+}$, 634.74 $[M+3H]^{3+}$.

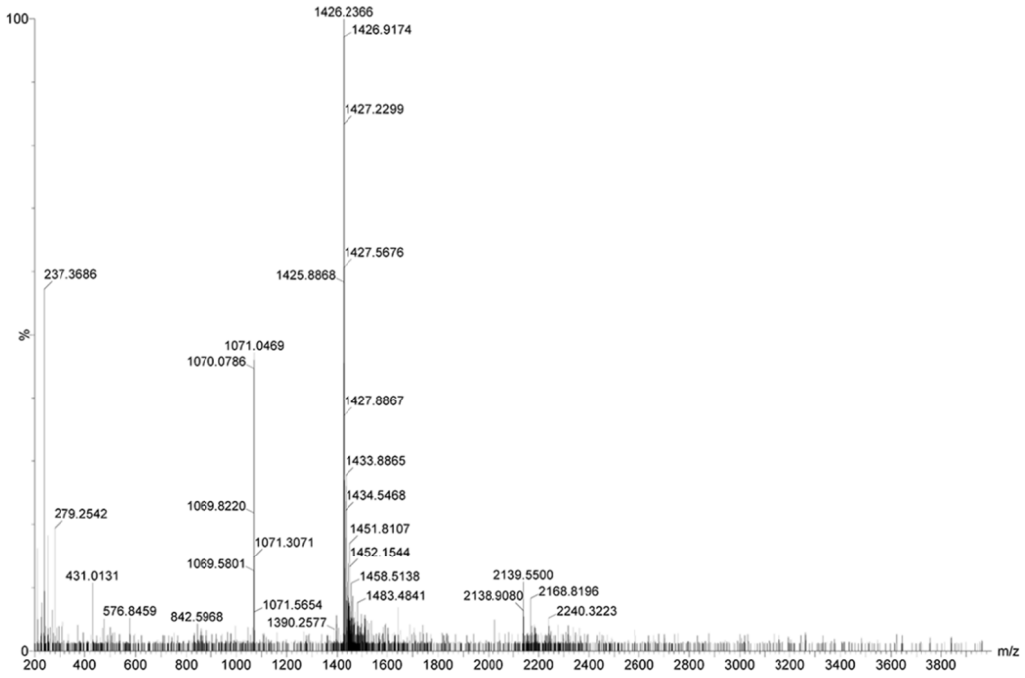


Figure S2.19. Mass spectrum of $(\mathbf{2})_9$ (retention time 12.6 min in Fig. 2.7c) from the LC-MS analysis of a DCL made from **2**. Calculated m/z: 2138.84 $[M+2H]^{2+}$, 1426.23 $[M+3H]^{3+}$, 1069.92 $[M+4H]^{4+}$; observed m/z: 2139.55 $[M+2H]^{2+}$, 1426.24 $[M+3H]^{3+}$, 1071.05 $[M+4H]^{4+}$.

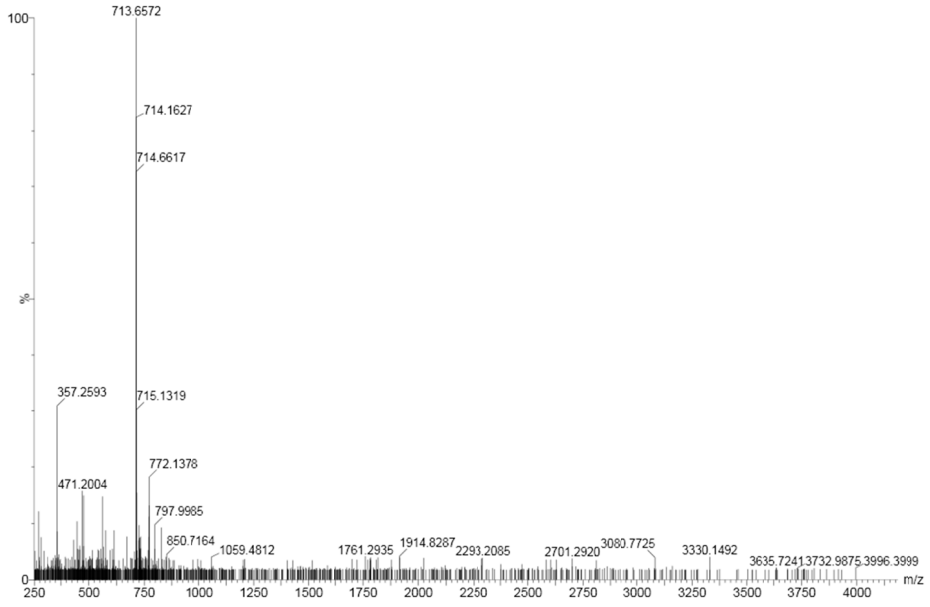


Figure S2.20. Mass spectrum of $(\mathbf{2})_3$ (retention time 12.9 min in Fig. 2.7c) from the LC-MS analysis of a DCL made from **2**. Calculated m/z: 713.62 $[M+2H]^{2+}$; observed m/z: 713.66 $[M+2H]^{2+}$.

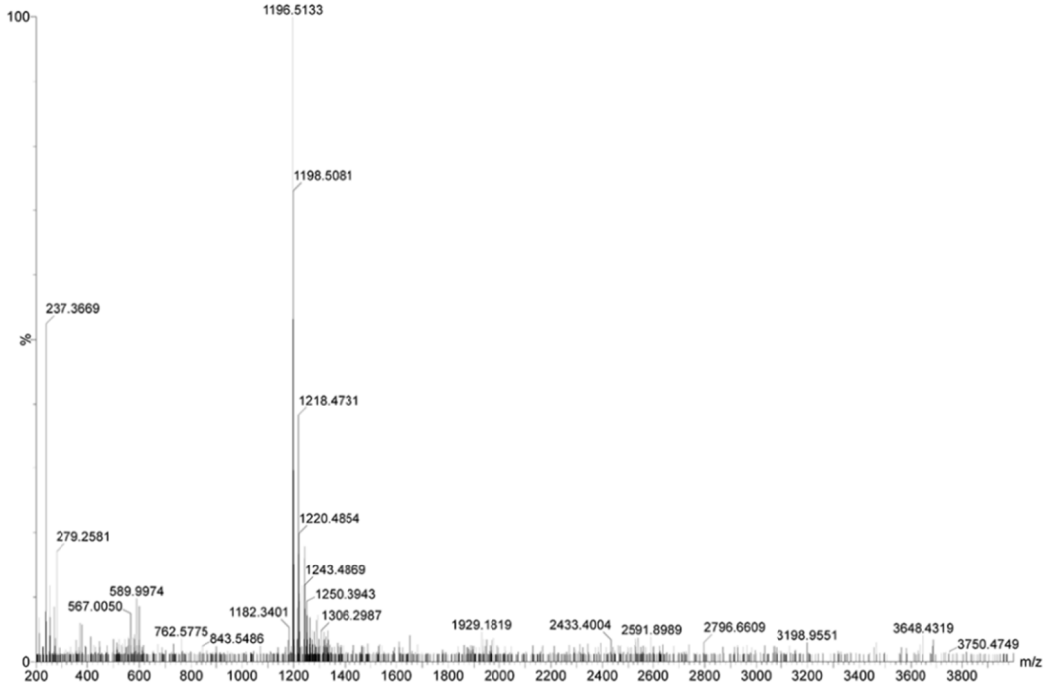


Figure S2.21. Mass spectrum of $(\mathbf{3})_4$ (retention time 6.3 min in Fig. 2.7d) from the LC-MS analysis of a DCL made from **3**. Calculated m/z: 1196.98 $[M+1H]^+$; observed m/z: 1196.51 $[M+H]^+$.

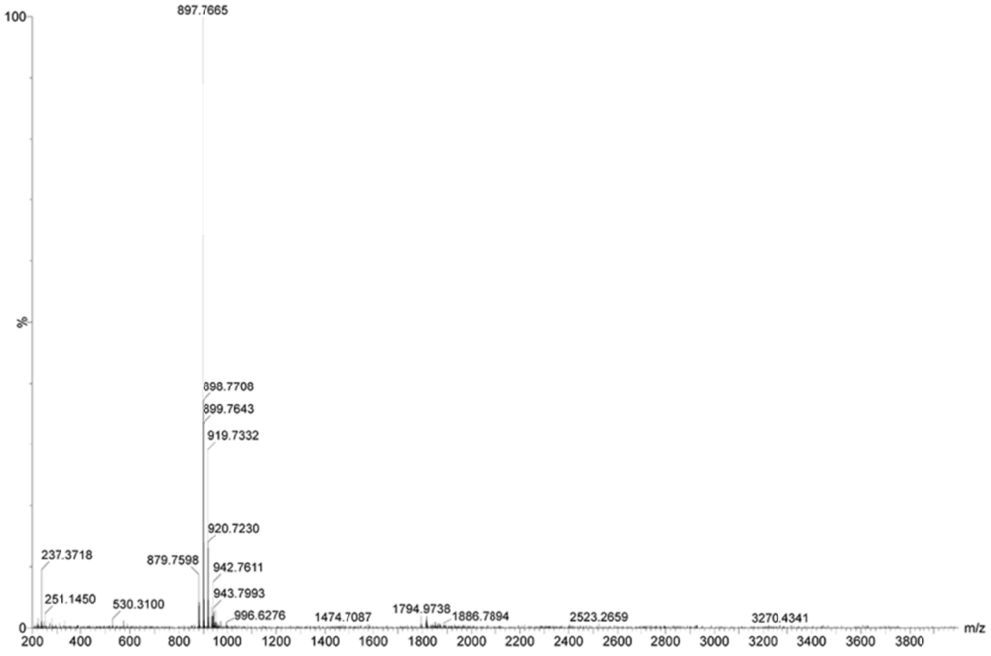


Figure S2.22. Mass spectrum of $(\mathbf{3})_3$ (retention time 7.6 min in Fig. 2.7d) from the LC-MS analysis of a DCL made from **3**. Calculated m/z: 897.98 $[M+1H]^+$; observed m/z: 897.77 $[M+1H]^+$.

2.7 References

(1) Lokey, R. S.; Iverson, B. L. *Nature* **1995**, *375*, 303.
(2) Hamuro, Y.; Geib, S. J.; Hamilton, A. D. *J. Am. Chem. Soc.* **1996**, *118*, 7529.
(3) Nelson, J. C.; Saven, J. G.; Moore, J. S.; Wolynes, P. G. *Science* **1997**, *277*, 1793.
(4) Bassani, D. M.; Lehn, J. M.; Baum, G.; Fenske, D. *Angew. Chem. Int. Ed.* **1997**, *36*, 1845.
(5) Appella, D. H.; Christianson, L. A.; Klein, D. A.; Powell, D. R.; Huang, X. L.; Barchi, J. J.; Gellman, S. H. *Nature* **1997**, *387*, 381.
(6) Berl, V.; Huc, I.; Khoury, R. G.; Krische, M. J.; Lehn, J. M. *Nature* **2000**, *407*, 720.
(7) Hill, D. J.; Mio, M. J.; Prince, R. B.; Hughes, T. S.; Moore, J. S. *Chem. Rev.* **2001**, *101*, 3893.
(8) Goodman, C. M.; Choi, S.; Shandler, S.; DeGrado, W. F. *Nat. Chem. Biol.* **2007**, *3*, 252.
(9) Horne, W. S.; Gellman, S. H. *Acc. Chem. Res.* **2008**, *41*, 1399.
(10) Saraogi, I.; Hamilton, A. D. *Chem. Soc. Rev.* **2009**, *38*, 1726.
(11) Guichard, G.; Huc, I. *Chem. Commun.* **2011**, *47*, 5933.
(12) Zhang, D. W.; Zhao, X.; Hou, J. L.; Li, Z. T. *Chem. Rev.* **2012**, *112*, 5271.
(13) Le Bailly, B. A. F.; Clayden, J. *Chem. Commun.* **2016**, *52*, 4852.
(14) Muller, M. M.; Windsor, M. A.; Pomerantz, W. C.; Gellman, S. H.; Hilvert, D. *Angew. Chem. Int. Ed.* **2009**, *48*, 922.
(15) Chandramouli, N.; Ferrand, Y.; Lautrette, G.; Kauffmann, B.; Mackereth, C. D.; Laguerre, M.; Dubreuil, D.; Huc, I. *Nat. Chem.* **2015**, *7*, 334.
(16) Becart, D.; Diemer, V.; Salaun, A.; Oiarbide, M.; Nelli, Y. R.; Kauffmann, B.; Fischer, L.; Palomo, C.; Guichard, G. *J. Am. Chem. Soc.* **2017**, *139*, 12524.
(17) Fletcher, J. M.; Harniman, R. L.; Barnes, F. R. H.; Boyle, A. L.; Collins, A.; Mantell, J.; Sharp, T. H.; Antognozzi, M.; Booth, P. J.; Linden, N.; Miles, M. J.; Sessions, R. B.; Verkade, P.; Woolfson, D. N. *Science* **2013**, *340*, 595.
(18) Hosseinzadeh, P.; Bhardwaj, G.; Mulligan, V. K.; Shortridge, M. D.; Craven, T. W.; Pardo-Avila, F.; Retti, S. A.; Kim, D. E.; Silva, D. A.; Ibrahim, Y. M.; Webb, I. K.; Cort, J. R.; Adkins, J. N.; Varani, G.; Baker, D. *Science* **2017**, *358*, 1461.
(19) Nair, R. V.; Vijayadas, K. N.; Roy, A.; Sanjayan, G. J. *Eur. J. Org. Chem.* **2014**, 7763.
(20) Li, J. W.; Nowak, P.; Otto, S. *J. Am. Chem. Soc.* **2013**, *135*, 9222.
(21) Corbett, P. T.; Leclaire, J.; Vial, L.; West, K. R.; Wietor, J. L.; Sanders, J. K. M.; Otto, S. *Chem. Rev.* **2006**, *106*, 3652.
(22) Case, M. A.; McLendon, G. L. *J. Am. Chem. Soc.* **2000**, *122*, 8089.
(23) Oh, K.; Jeong, K. S.; Moore, J. S. *Nature* **2001**, *414*, 889.
(24) Krishnan-Ghosh, Y.; Balasubramanian, S. *Angew. Chem. Int. Ed.* **2003**, *42*, 2171.
(25) Woll, M. G.; Gellman, S. H. *J. Am. Chem. Soc.* **2004**, *126*, 11172.
(26) Roy, L.; Case, M. A. *J. Am. Chem. Soc.* **2010**, *132*, 8894.
(27) Lao, L. L.; Schmitt, J. L.; Lehn, J. M. *Chem.- Eur. J.* **2010**, *16*, 4903.
(28) Lafuente, M.; Atcher, J.; Sola, J.; Alfonso, I. *Chem.- Eur. J.* **2015**, *21*, 17002.
(29) Tsiamantas, C.; de Hatten, X.; Douat, C.; Kauffmann, B.; Maurizot, V.; Ihara, H.; Takafuji, M.; Metzler-Nolte, N.; Huc, I. *Angew. Chem. Int. Ed.* **2016**, *55*, 6848.
(30) Otto, S.; Furlan, R. L. E.; Sanders, J. K. M. *Science* **2002**, *297*, 590.
(31) Lam, R. T. S.; Belenguer, A.; Roberts, S. L.; Naumann, C.; Jarrosson, T.; Otto, S.; Sanders, J. K. M. *Science* **2005**, *308*, 667.
(32) Ponnuswamy, N.; Cougnon, F. B. L.; Clough, J. M.; Pantos, G. D.; Sanders, J. K. M. *Science* **2012**, *338*, 783.
(33) Carnall, J. M. A.; Waudby, C. A.; Belenguer, A. M.; Stuart, M. C. A.; Peyralans, J. J. P.; Otto, S. *Science* **2010**, *327*, 1502.
(34) Frederix, P. W. J. M.; Ide, J.; Altay, Y.; Schaeffer, G.; Suring, M.; Beljonne, D.; Bondarenko, A. S.; Jansen, T. L. C.; Otto, S.; Marrink, S. J. *ACS Nano* **2017**, *11*, 7858.

(35) O'Sullivan, M. C.; Sprafke, J. K.; Kondratuk, D. V.; Rinfray, C.; Claridge, T. D. W.; Saywell, A.; Blunt, M. O.; O'Shea, J. N.; Beton, P. H.; Malfois, M.; Anderson, H. L. *Nature* **2011**, *469*, 72.
(36) Claessens, C. G.; Stoddart, J. F. *J. Phys. Org. Chem.* **1997**, *10*, 254.
(37) De, S.; Chi, B.; Granier, T.; Qi, T.; Maurizot, V.; Huc, I. *Nat. Chem.* **2018**, *10*, 51.
(38) Kabsch, W. *Acta Crystallogr D Biol Crystallogr* **2010**, *66*, 133.
(39) Uson, I.; Sheldrick, G. M. *Acta Crystallogr D* **2018**, *74*, 106.
(40) Sheldrick, G. M. *Acta Crystallogr. Sect. A: Found. Crystallogr.* **2008**, *64*, 112.
(41) Brandenburg, K.; Berndt, M. *Crystal Impact GbR, Bonn, Germany* **1999**.